9-2009

Distribution and Diversity of Planktonic Ciliates: Patterns and Processes

Mary Doherty
University of Massachusetts Amherst, themaryd@gmail.com

DISTRIBUTION AND DIVERSITY OF PLANKTONIC CILIATES: PATTERNS AND PROCESSES

A Dissertation Presented

by

MARY DOHERTY

Submitted to the Graduate School of the University of Massachusetts Amherst in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2009

Program in Organismic and Evolutionary Biology

DISTRIBUTION AND DIVERSITY OF PLANKTONIC CILIATES: PATTERNS AND PROCESSES

A Dissertation Presented

by

MARY DOHERTY

Approved as to style and content by:

Laura A. Katz, Chair

Benjamin B. Normark, member

George B. McManus, member

Rob Dorit, member

Joseph S. Elkinton, Director
Organismic and Evolutionary Biology

DEDICATION

To my daughter, Ruth A. Doherty, keeping in mind that what is essential is invisible to the eye.

ACKNOWLEDGEMENTS

I would like to acknowledge my thesis advisor, Laura A. Katz, who has inspired my science, provided me with all the tools I needed to succeed, told me what I needed to hear, no matter how unpleasant, and demonstrated to me how to be a genuine advocate for students. I would also like to acknowledge the members of the Katz lab past and present who gave endless support scientifically and personally, and whose friendship I will treasure, particularly Micah Dunthorn, Rebecca Zufall, and Laura Parfrey.

If it takes a village to raise a child, it takes a small city to support a single mother and her daughter through a Ph. D. program. So many people gave of themselves unconditionally to make this happen, particularly, my mother, Anne Doherty, my brother, George Doherty, and my friends Lisa Wall, Wendy Zucco, Amy Montali, and Susan McCarthy. I am endlessly grateful to them and many others who helped me along through the hardest times.

ABSTRACT

DISTRIBUTION AND DIVERSITY OF PLANKTONIC CILIATES:

PATTERNS AND PROCESSES

SEPTEMBER 2009

MARY DOHERTY, B.A., SMITH COLLEGE

M. Ed., UNIVERSITY OF MASSACHUSETTS AMHERST

M. A., SMITH COLLEGE

Ph. D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Laura A. Katz

The nature and extent of microbial biodiversity remain controversial with persistent debates over patterns of distributions (*i.e.* cosmopolitanism vs. endemism) and the processes that structure these patterns (neutrality vs. selection). I used culture-independent approaches to address these issues focusing on two groups of ciliates, the Oligotrichia (Spirotrichea) and Choreotrichea (Spirotrichea). To assess the diversity of these ciliates, I designed primers specific to SSU rDNA of ciliates within these clades, and investigated (1) geographic and temporal distributions along three coastal sites in the Northwest Atlantic; (2) the relationship between ciliate communities in the benthos and the plankton along the New England coast; and (3) diversity in ciliate communities across an environmental gradient at six stations in Long Island Sound spanning the frontal region that separates the fresher Connecticut River outflow plume from the open Sound.

Each collection had its own distinct assemblage of rare and abundant ciliate haplotypes, and genealogical analyses of my samples combined with published sequences

from identified morphospecies reveal that haplotype diversity at these sites is greatest within the genus *Strombidium*, in the Oligotrichia. Clustering of phylogenetic types indicates that benthic assemblages of oligotrichs and choreotrichs appear to be more like those from spatially distinct benthic communities than the ciliate communities sampled in the water above them. Neither ciliate diversity nor species composition showed any clear relationship to measured environmental parameters (temperature, salinity, accessory pigment composition, and chorophyll), although I observed that diversity decreased moving from nearshore to offshore. I find no strong fit of my communities to log series, geometric, or log normal distributions, though one of the 3 clusters is most consistent with a log series distribution. These analyses suggest that Oligotrich and Choreotrich communities in coastal environments may be distributed in a neutral manner.

I investigated the effectiveness of molecular approaches in characterizing ciliate diversity in my samples. Estimates of diversity based on molecular markers are similar to estimates from morphological observations for Choreotrich ciliates, but much greater for Oligotrich ciliates. Sediment and plankton subsamples differed in their robustness to repeated subsampling. Sediment gave variable estimates of diversity while plankton subsamples produced consistent results.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

CULTURE-INDEPENDENT ASSESSMENT OF PLANKTONIC CILIATE DIVERSITY IN COASTAL NORTHWEST ATLANTIC WATERS

Introduction

Elucidating patterns of ciliate diversity in marine systems is essential because ciliates play key roles in marine food webs. Ciliates are a trophic link between nano- and picoplankton, and larger metazoan organisms (Pierce and Turner 1992, Calbet and Saiz 2005), and are important grazers on harmful algal bloom species (Rosetta and McManus 2003, Kamiyama and Matsuyama 2005). Microbial assemblages shift and change with biotic and abiotic processes (Fuhrman *et al.* 2006), and hence identifying the abundance and diversity of key organisms is essential to understanding ecosystem functions.

This study combines molecular methods and microscopy to elucidate the diversity of two ecologically important subclasses of ciliates – Oligotrichia and Choreotrichia. These two morphologically diverse subclasses dominate ciliate communities in near-coastal waters. For example, in a three-year study in Long Island Sound, (Capriulo and Carpenter 1983, Capriulo *et al.* 2002) reported the presence of 65 species from these groups based on light microscopy, including 30 species of tintinnids, 12 other choreotrich species, and 23 oligotrichs.

The use of morphological characters alone to assess ciliate community diversity is problematic in ecological studies, since for many ciliates identification to species is level is possible only using fine structure. For example, some ciliate groups abundant in the plankton can only be distinguished to coarse morphological categories, such as “small aloricate ciliates” (Sherr *et al.* 1986, Fileman and Leakey 2005), or as highly diverse

assemblages within a genus, such as *Strombidium spp.* (Modigh 2001, Fileman and Leakey 2005). One exception is the tintinnids, a group of choreotrichs with a rich literature of morphological species descriptions based mainly on size and shape of the lorica (outer sheath). Ocean transect studies of tintinnid diversity reveal distinct assemblages within ocean provinces, with specific diversity patterns characterizing each area (Modigh *et al.* 2003, Thompson 2004). Further, compiling 451 data points from the literature for tintinnid diversity, Dolan *et al.* (2006) found a latitudinal distribution of tintinnid richness.

Contemporary views on diversity of marine microbes have been altered by large-scale, molecular surveys (Caron *et al.* 2004, DeLong 2005, Giovannoni and Stingl 2005, Richards and Bass 2005, Xu 2006). Such studies revealed considerable microbial diversity not captured by culture-dependent methods. With these new methods, it is possible to sample from a broader range of microbial habitats (Dawson and Pace 2002, Edgcomb *et al.* 2002, Zettler *et al.* 2002, Lopez-Garcia *et al.* 2003), to sample previously undetected microbes (Diez *et al.* 2001, Moon-van der Staay *et al.* 2001), and to determine the spatial and temporal scales of microbial diversity (Romari and Vaulot 2004, Behnke *et al.* 2006, Cordova-Kreylos *et al.* 2006, Ley *et al.* 2006). Although ciliates are often captured in studies of eukaryotic diversity, I know of no published studies focusing specifically on ciliate diversity using culture-independent molecular methods.

I combined molecular methods and microscopy to identify ciliates in the subclasses Oligotrichia and Choreotrichia across two seasons – fall and spring – at three geographically distinct coastal sites in the Northwest Atlantic. In addition to evaluating

morphological and genetic diversity, I measured a few key environmental parameters that may influence ciliate distributions.

Materials and Methods

Ciliates were sampled from three near shore locations: two in the Gulf of Maine, and one in Long Island Sound (LIS) (Fig 1.1; Table 1.1). Samples from Maine and Connecticut were collected on the same day for two time points, one in October, 2004 and one in May, 2005. For molecular analysis, I collected 50-60L of surface waters per sample and concentrated the microplankton down to 1L by siphoning through a 20µm mesh. Of this 1L, five 200ml aliquots were filtered onto a 5µm Millipore cellulose nitrate filter. Each filter was placed in 1ml of DNA prep buffer (100mM NaCl, Tris-EDTA at pH 8, and 0.5%SDS) until DNA extraction. For microscopic examination, 250 ml surface water samples were preserved in 5% Lugols iodine solution.

I recorded ambient temperature, and salinity at each collection. For the May 2005 samples, I also measured chlorophyll concentration. For chlorophyll, 100 ml of each pre-concentrated water sample was filtered onto a Whatman GFF glass fiber filter. The filter was folded in half, wrapped in aluminum foil and placed on dry ice, and later stored at -80°C prior to extraction in 90% aqueous acetone and quantification by fluorescence.

Total genomic DNA was extracted from the samples using standard protocols (Ausubel *et al.* 1992). Two of the samples (Southport Island, October and Connecticut, May) were difficult to amplify, so I extracted these filters with the DNeasy plant kit by Qiagen (cat. # 69104) to remove compounds inhibitory to PCR.

To design clade-specific primers, I searched GenBank for all full-length SSU rDNA sequences from the ciliate class Spirotrichea. I aligned a total of 69 sequences

representing all five subclasses: Oligotrichia, Choreotrichia, Hypotrichia, Stichotrichia, and Protocruziidia using the ClustalW algorithm in the program MegAlign (DNAStar Inc.). I then determined optimal sites for primer design by visual inspection in MacClade version 4.06. My aim was to choose primers in regions that are conserved within the Choreotrichia and Oligotrichia, but divergent in the other subclasses. However, there were very few regions in the Spirotrichea that were specific to only Oligotrichia and Choreotrichia without also matching the Stichotrichia, so I allowed for some overlap in sequence with this clade. My rationale was that Stichotrichs are predominantly sediment dwelling and unlikely to dominate planktonic samples. The four primers differed from the nearest sister taxa (members of the Hypotrichia) by an average of 22.4% (4.75 bases) with a range of 15-33% (3-7 bases). To increase specificity, I designed the primers to concentrate the divergence at the 3' ends. I selected primer pairs from two distinct, non-overlapping regions of the SSU marker to compare efficiency in capturing ciliate community diversity using published sequences from identified morphospecies as the backbone for my analyses. The efficiency of these primer pairs was tested by amplifying genomic DNA of cultured specimens within the Choreotrichia and Oligotrichia, including *Metacylis angulata, Laboea strobila, Strombidium oculatum, Strombidium stylifer, Tintinnopsis sp., and Eutintinnus pectinis.*

I amplified DNA fragments using polymerase chain reaction (PCR) with Phusion polymerase and reagents from Finnzyme Inc. Each sample was amplified using both primer sets. The first primer pair (primer set A) consisted of 150+: 5' AHTTACATGGATAACCGTGG and 568-: 5' GGTSTAAATTCRKYTCATTKC. Cycling conditions for primer set A were as follows: 98° for 2 minutes, 35 cycles of 98°

for 15 seconds, 55° for 30 seconds, and 72 ° for 45 seconds, then a 10 minute extension at 72°. The second primer pair (primer set B) consisted of 1199+: 5' GCCGACTCGGGATCGGGGGC and 1765-: 5' CCCCAKCACGACDCMTATTGCTG. Cycling conditions for primer set B were as follows: 98° for 2 minutes, 35 cycles of 98° for 15 seconds, 72° for 1 minute 30 seconds, then a 10 minute extension at 72°. Three separate 20µl reactions were run and the products pooled to minimize PCR bias. Primer set A amplifies a 418 bp fragment of the SSU rDNA gene, and Primer set B amplifies a 566 bp region.

PCR products were gel-isolated, and cleaned using the UltraClean GelSpin DNA purification kit from Mo Bio Laboratories (cat. #12400-100). I used the pSTBlue-1 Perfectly Blunt Cloning Kit from Novagen (cat. # 70191-4DFRZ) for cloning, and then picked and miniprepped 192 colonies per sample using either the Plasmid 96 Miniprep Kit from Edge Biosystems (cat. # 49181) or the PureLink 96 Plasmid Purification System from Invitrogen (cat. # 12263-018). Sequencing reactions were performed using Big Dye Termination Kit (Applied Biosystems), cleaned with a sephadex plate column, and sequenced on an ABI 377 automated sequencer.

I assembled and edited sequences using SeqMan (DNAStar Inc.). I explored assembling my sequences at different similarity cutoffs in SeqMan, which allows assembly of haplotypes with varying parameters (*e.g.* allowing sequences to cluster that are within 1% of one another). Based on comparison of similarities set from 90-100%, I selected a 99% similarity cutoff for genealogical analyses and diversity estimation to allow for inclusion of highly related but distinct taxa. Haplotypes were then checked for identity with published sequences using BLAST search (Altschul *et al.* 1997) on NCBI.

For genealogical analyses, haplotypes were aligned with published sequences from identified morphospecies obtained by searching GenBank for all entries recorded as Choreotrichia and Oligotrichia. Although included in my BLAST assessments, I chose to exclude environmental samples from my genealogical analyses in order to focus on a taxonomic framework for interpreting my haplotypes. I used the CLUSTAL W algorithm as implemented in MegAlign to align my sequences with the published sequences. I finalized alignments by eye in MacClade version 4.06.

To check my assembled haplotype sequences for PCR artifacts such as chimeras I scanned my alignments in both the Chimaera (Maynard Smith 1992, Posada and Crandall 2001, Posada 2002) and GENECONV (Padidam *et al.* 1999) applications in RDP version 2.0, recombination detection software (Martin *et al.* 2005). I then visually inspected the aligned sequences in MacClade version 4.06 to detect recombination events. I detected recent recombination, or PCR chimerism, in only one haplotype, sequenced by primer set A, and removed it from the analysis.

Bayesian analyses were conducted for each primer dataset using a GTR + G + I model of sequence evolution in MrBayes (Ronquist and Huelsenbeck 2003). Four simultaneous MCMCMC chains were run for 3,000,000 generations sampling every 100 generations. Stationarity in likelihood scores was determined by plotting the -1nL against the generation. All trees below the observed stationarity level were discarded, resulting in a "burn in" of 75,000 generations. Estimation of best fit models for partial SSU rDNA gene sequences were performed using MrModelTest 2.2 (Nylander 2004).

I estimated the number of morphospecies in the sample by settling two aliquots of 100 ml each and examining the entire settled volume at 400-600x on an Olympus

inverted microscope. Morphospecies were documented by drawing and digital photography. Using this approach, most ciliates were recognizable to genus based on overall morphology. For the tintinnids, an extensive literature on species diagnosis is available, based on lorica morphology. Unfortunately, the lorica has been shown to be a plastic character within species (Alder 1999). For non-tintinnid choreotrichs, and for oligotrichs, species diagnosis is based on silver-staining procedures and/or electron microscopy, methods that are best applied when an abundance of specimens is available, for example from cultures (Montagnes and Lynn 1991) Lynn and Small 2000). Moreover, the fixative that provides the best quantitative preservation, Lugol's iodine solution, obscures many cytological features. For these reasons, I used simple light microscopy and examined my samples conservatively, identifying separate morphospecies only when clear differences in morphology were present. Principally, I used the keys in (Marshall 1969) for tintinnids and (Maeda and Carey 1985, Maeda 1986) for other choreotrichs and for oligotrichs.

To compare species diversity between molecular samples, rarefaction curves were calculated using EstimateS version 7.5 (Colwell 2005), comparing number of clones sequenced to number of observed haplotypes based on my 99% assembling criterion. I also calculated the non-parametric richness estimator, Chao1, with EstimateS using 100 randomizations, sampling without replacement.

Results

My clone libraries generated a rich diversity of haplotypes within the Choreotrichia and Oligotrichia. For brevity, I refer to the site names in the study as follows: Bucks Harbor, ME (BH), Southport Island, ME (SI), and Groton, CT (CT)

(Table 1.1). The clade-specific primers were successful in amplifying the target groups of Oligotrichia and Choreotrichia, and excluding non-target groups, as evidenced in my phylogenies (Figs. 1.2 and 1.3). The few exceptions to this were in the sister taxa, Stichotrichia (Fig. 1.3). In spite of the degeneracy in the primer sequences, my phylogenies and BLAST search results show no evidence of amplification of non-ciliates, or even of non-Spirotrichs (Figs. 1.2 and 1.3).

I sequenced a total of 731 clones with primer set A, and 653 clones with primer set B (Table 1.2). To explore the patterns of divergence among the resulting sequences, I assessed the number of haplotypes generated using a range of similarity cutoffs at 15 intervals, collapsing sequences that were 90-100% identical into single haplotypes (see methods). Analysis of the impact of this range of bins with the two primer sets revealed a large increase in the number of haplotypes between 97-99% similarity (Fig. 1.4), with minimal change when increased to 100%. In other words, there are few clones that differ from one another by ≤1%, and this level of diversity includes experimental error. Thus, for the purposes of comparing sites and generating genealogies, I used the 99% assembly criterion.

Based on an assembly of clones that were ≥99% similar, 56 haplotypes and 66 haplotypes were generated from primer sets A and B, respectively (Table 1.2) (GenBank Accession Numbers EF553335–EF553457). Of these haplotypes, I found only one (hbp23, amplified with primer set B) that occurred at all sites and times, and a small proportion (13 of 123) of the haplotypes were sampled in three or more of the six collections. The bulk of the haplotypes I sequenced with both primer sets (110 of 123)

were rare, occurring in only one or two samples. This rarity was consistent across both primer sets, representing 89% of the haplotypes in both cases.

Four of the 57 haplotypes obtained with primer set A were 100% identical to published sequences as determined by BLAST (Altschul *et al.* 1997): hap40 = *Pelagostrobilidium neptuni* (GB AY541683), hap44 = *Strombidinopsis jeokjo* (AJ628250), hap27 = *Laboea strobila* (AF399153), and hap36 = *Strombidium biarmatum* (AY541684) (Table 1.3). Four of the 66 haplotypes sequenced with primer set B were 100% identical to published sequences: hbp65 = *Tintinnopsis tubulosoides* (AF399111), hbp71 = an uncultured eukaryote (AY129053), hbp20 = *Strombidium biarmatum* (AY541684) and hbp23 = *Pelagostrobilidium neptuni* (AY541683) (Table 1.3). The latter two, *Strombidium biarmatum* and *Pelagostrobilidium neptuni* were captured with both primer sets A and B.

I sampled the greatest diversity within the Oligotrichia (Figs. 1.2 and 1.3). Primer sets A and B sequenced 38 of 57 and 42 of 66 haplotypes that fell within this clade, respectively. My most common and abundant haplotype is identical to published sequences of *Strombidium biarmatum*. This haplotype sequence was sampled 184 times with primer set A, and 251 times with primer set B (Table 1.3) and falls within the Oligotrichia clade as hap36 (Fig. 1.2) and hbp23 (Fig. 1.3), respectively. Within the Choreotrichia, the two primer sets captured similar levels of diversity. Primer set A sampled 19 of 57 haplotypes within this clade, and Primer set B sampled 21 of 66 haplotypes. I detected non-target sequences from the closely related group, Stichotrichia, in three of my samples using primer set B. These included a haplotype 98% similar to *Holosticha diademata* (DQ059583) in the Bucks Harbor October sample, and a haplotype

96% similar to *Gonostomum strenuum* (AJ310493) in the Bucks Harbor May and Connecticut October samples (Fig 1.3).

Two of the samples I collected, SI Oct and CT May were difficult to amplify with PCR, and required a plant-specific extraction protocol to obtain genomic DNA suitable for PCR (see methods). The two samples were much lower in haplotype diversity than any of the others when amplified using primer set A, but not with primer set B (Table 1.4). In addition, morphological and molecular results are inconsistent in the CT May sample, which was quantified using both methods (Table 1.5). Choreotrichs that were present in the sample did not amplify, while numbers of Oligotrichs similar to the other samples did amplify. These data suggest the water sampled at the CT May and the SI Oct sites contained a component inhibitory to PCR and/or that the community composition in these samples contain fewer members, as seen in other studies (Costas *et al.* 2007).

I calculated rarefaction curves, again based on haplotypes defined to be 99% similar (Figs. 1.7 and 1.8) using EstimateS version 7.5 (Colwell 2005). Rarefaction calculates the number of haplotypes as a function of the number of sequences sampled, and is a means of comparing diversity across different sample sizes. I estimated diversity within each sample collected and plotted the curves for each on the same graph to compare haplotype richness between samples (Figs. 1.7 and 1.8). The trajectories of the curves indicate that I have not sampled all of the diversity for the collection sites, as the slopes of the lines have not reached an asymptote (Figs. 1.7 and 1.8). Comparing diversity between collections, the 95% confidence intervals overlap, indicating no significant differences in diversity as estimated by rarefaction. There is a trend toward higher diversity in the SI May sample based on the rarefaction trajectories for both primer

sets A and B (Figs. 1.7 and 1.8). The SI Oct and CT May samples sequenced with primer set A have lower diversity, having a shallower slope than the other curves. These were the samples that were difficult to amplify.

I used the Chao1 diversity estimator to estimate total haplotype diversity of the samples (Table 1.4). This is a non-parametric estimator of the total diversity in the community from which the sample is drawn. Using one hundred randomizations without replacement, I estimated diversity against number of clones sampled for each collection separately for the two primer sets, and reported the total estimated diversity (Table 1.4). The data from the Chao1 estimator is consistent with my rarefaction data. SI May had the highest estimates of total diversity, 325 and 253 haplotypes for primers A and B respectively (Table 1.4). My results for the SI Oct and CT May samples were inconsistent across primer sets A and B, reflected in the Chao1 diversity estimates. Based on results from primer set A, the Chao1 estimate of total diversity was only 6 haplotypes for SI Oct, and 28 for CT May (Table 1.4). Using the data from primer set B for each of these samples, SI Oct was estimated to contain 136 haplotypes and CT May was estimated to contain 78 haplotypes (Table 1.4). Estimated sample diversity was not correlated with the abiotic factors I measured (temperature, salinity, latitude and season; Table 1). However, there is a trend of higher haplotype diversity estimated by rarefaction and Chao1 (Figs. 1.7 and 1.8: Table 1.4) associated with high chlorophyll content in the SI May sample.

As is typical for spring in LIS, ciliates from the CT May sample were abundant in the plankton (>2000 cells l^{-1}). I observed 16-19 different ciliate morphospecies in 200 ml total of settled sample. The majority of these were from the Spirotrich subclasses

Oligotrichia and Choreotrichia (the latter including the Tintinnids; Table 1.5). The only ciliate that was identified to the species level was *Laboea strobila*, easily recognizable by its spirally-wound girdle and large size (McManus and Fuhrman 1986). This mixotrophic oligotrich is common in LIS during spring and early summer. My molecular approach did not capture *Laboea strobila* at this sampling location and time (Figs. 1.2 and 1.3). There were at least four other distinct morphospecies within the same family (Strombidiidae), including probably *Strombidium conicum* (Table 1.5).

Microscopic observations suggested that the spring ciliate assemblage in the plankton of coastal Maine was similar to that sampled simultaneously from Long Island Sound. *Laboea strobila* was present in both samples, with higher abundances at Southport Island (c. 300 cells l^{-1}). Molecular samples are consistent with these observations in both Maine locations. Southport Island clone libraries also contain *Laboea strobila* in high abundances (Figs. 1.2 and 1.3). Other Strombidiids could be classified into 4-5 morphospecies. Most of the Tintinnids at both sites were from the agglutinated genus *Tintinnopsis*. One empty lorica of the hyaline genus *Parafavella* (probably *P. denticulata*) was observed in the Southport Island sample (Table 1.5). My molecular genealogies show that in Maine I sampled haplotypes closely related to published *Tintinnopsis* and *Parfavella* genera (Figs. 1.2 and 1.3).

Discussion

I observed similar overall levels of diversity across three sites and two seasons, with only a few haplotypes found multiple times in clone libraries. The limitations inherent in using PCR of environmental samples to assess abundance and diversity are well known and I recognize that, in the absence of comparisons with quantitative PCR

results, assessments of frequencies in clone libraries are at best a rough approximation of abundance in nature. The potentially differing levels of amplification in ciliate macronuclear genomes further exacerbate this problem. I attempted to reduce the bias in my amplification methods by running multiple PCR reactions and pooling the product pairs in generating clone libraries. In addition, I tested the cloning results with one of my samples by sequencing twice the number of clones, and found no difference in the diversity estimates with additional sampling. Moreover, I found that my clade-specific primers were consistent in the levels of diversity they identified, and that my methods were reproducible, in that when I repeated the PCR, cloning and sequencing methods for two of the samples, I saw no major differences between trials.

Although the level of diversity was similar across sites (with the possible exception of the elevated diversity at the chlorophyll-rich SI Oct site), there is a unique assemblage of haplotypes at each sampling location and time (Figs. 1.5 and 1.6). This contrasts with some studies based on morphological observations of overall ciliate diversity, which reveal a striking homogeneity at local and global scales of ciliate species assemblages (Finlay 2002). My findings are consistent with the studies on tintinnids (Modigh *et al.* 2003, Thompson 2004, Dolan *et al.* 2006), where specific patterns of diversity are associated with regional areas in the ocean. The observation in my study is that morphospecies underestimate genetic diversity. Nevertheless, similarity between my molecular results and the tintinnid data indicate that a pattern of varying assemblages may prevail within planktonic ciliates.

The estimates of overall diversity vary between the molecular and morphological sampling methods. Based on microscopic observation using standard ecological

techniques, I conservatively estimated that 12-15 morphospecies from the Choreotrichia and Oligotrichia were present in each sample. Based on my molecular data, Chao1 diversity calculated with EstimateS version 7.5 (Colwell 2005) for each set of haplotypes sampled is substantially higher (Table 1.4). The average estimated diversity of haplotypes per collection is 133, an approximately 100-fold difference from the morphospecies estimate. Among the possible explanations for the apparent underestimate of morphological methods include: (1) cryptic species of ciliates; (2) high intraspecific variation within morphospecies and (3) undercounting due to lower volume sampled for the time-intensive morphological estimates. Distinguishing among these hypotheses must await further molecular and genetic analyses of morphospecies.

When I look within the subclasses, I find that diversity estimates based morphological observations are consistent with haplotype frequencies in the Choreotrichia, but not in the Oligotrichia. For my two Maine sites, I observed approximately the same number of Tintinnid and other Choreotrich morphospecies as haplotypes sequenced (Table 1.5). The Connecticut May sample, which was difficult to amplify, is not consistent with this trend, but I predict that this is a result of PCR inhibition in that specific sample (Table 1.5). In contrast with the Choreotrichia, the number of haplotypes sequenced in the Oligotrichia, particularly those falling within the *Strombidium* genus on my phylogeny (Figs. 1.2 and 1.3), was 2-3 times greater than the number of distinct morphospecies I observed microscopically from this clade (Table 1.5). The greater diversity of genetic versus morphological entities within my samples suggests that there may be cryptic species of ciliates that are not readily distinguishable

by light microscopy in the Oligotrichia, or that there is an as yet unappreciated diversity of very small ciliates, which are difficult to quantify by microscopy (Sherr *et al.* 1986).

To assess the taxonomic position of haplotypes, I generated genealogies including published sequences from identified morphospecies and excluded unknown environmental samples (Figs. 1.2 and 1.3). Notably, the bulk of the diversity I observed is within the *Strombidium spp.* clade, which yielded 37 (primer set A) and 42 (primer set B) haplotypes (Figs. 1.2 and 1.3). Given that a total of 63 morphospecies have been described in this clade worldwide (Agatha 2004), I would not have predicted such high numbers of haplotypes falling within this clade purely from marine planktonic samples. The most abundant and widespread haplotype I sampled was 100% identical to *Strombidium biarmatum*, found in high numbers and in almost every sample. If the genetic diversity I obtained is reflective of species diversity, the number of haplotypes I obtained in the *Strombidium* clade may indicate high levels ecological complexity within marine microenvironments, where many similar species are able to coexist by fine-scale partitioning of niches. Alternatively, these data may reflect high levels of standing genetic variation within a small number of ciliate species. Data on other planktonic ciliates within these subclasses are not consistent with this second assertion (Katz *et al.* 2005), as I find extremely low levels of genetic variation (<0.5%) in *Laboea strobila* populations sampled over space and time.

Although the relationships between genetic and morphological species of ciliates are unclear, a 1% cutoff has been argued by some to be informative about species diversity for eukaryotic microbes for full-length SSU rDNA sequences (Richards *et al.* 2005, Stoeck *et al.* 2006). Using this 99% similarity cutoff between sequences as a

criterion for defining a haplotype, I find many more genetic entities than morphological entities in comparable samples. Notably, for Choreotrichia and Oligotrichia, the levels of genetic variation in the two regions assessed here were comparable to divergences for the full-length sequence. To assess the relevance of a 1% cutoff, I tabulated distance between data from species and genera. Pairwise genetic distances between SSU rDNA sequences from published morphospecies within Choreotrichia and Oligotrichia vary widely, ranging from 0-3.7% divergence between species, while distances between genera range from 1.7-8.7% divergence.

Comparisons of diversity within Oligotrichia by assembling at a lower percentage of similarity indicate that the number of haplotypes does not drop to the levels estimated by morphology until I get to <95% similarity. Prior studies combining morphological and molecular data also indicate high genetic variation within morphologically similar members of the Oligotrichia (Snoeyenbos-West *et al.* 2002, Katz *et al.* 2005).

Haplotype diversity in my samples is consistent with environmental samples of marine eukaryotes in general (Countway *et al.* 2005). Universal eukaryotic primers detected high numbers of rare haplotypes, and relatively few abundant haplotypes, and were able to detect changing community compositions within a sample. Recent analyses of prokaryotic diversity in the coastal Pacific Ocean revealed predictable patterns of change in community composition over seasonal and interannual scales (Fuhrman *et al.* 2006). Thus, the genetic variation I see in ciliate assemblages may indicate adaptations to diverse and variable microbial communities in lower trophic levels.

By sampling with primers designed to amplify two different regions of the SSU rDNA locus, I identified differences within a single sample that may be the result of

preferential amplification of DNA templates. However, my phylogenetic data and my diversity estimates were comparable across primer sets, containing similar numbers of haplotypes within the higher taxonomic groupings of Choreotrichia and Oligotrichia (Figs. 1.2 and 1.3). Ultimately, this preliminary work can be applied to additional studies designing FISH probes, or amplification of samples using the forward primer of set A and the reverse primer of set B to get a larger sequence.

Using clade-specific primers and focusing on these two groups of ciliates has allowed us to sample these coastal environments more thoroughly than previous environmental studies. My data reveal complex diversity patterns across time and space. In addition, I find that molecular tools enabled us to sample diversity I did not find using morphological approaches. In the Oligotrichia, I observed no greater than 6 morphospecies total, compared to the 38-40 haplotypes I sampled with primer sets A and B (Table 1.5). I can use these to look for patterns and driving forces of ciliate diversity, and for morphological attributes that correlate with the molecular diversity.

Table 1.1 Sampling data for all collection sites and time points showing coordinates, temperature, salinity, and chlorophyll concentration.

Location	Abbr.	Geographic Coordinates	Sampling Dates	Temp. ° C	Salinity ppt.	Chlorophyll (µg/L)
Bucks Harbor, ME	BH	44°38.20'N, 67°22.29'W	October, 2004	11.1	33	NA
			May 2005	6.7	30	2.0
Southport Island, ME	SI	43°49.05'N 69°39.16'W	October, 2004	13.3	32	NA
			May 2005	8.9	27.	10.4
Groton, CT	CT	41°19.00'N 72°03.65'W	October, 2004	18.8	29	NA
			May 2005	12.1	25	1.4

Table 1.2 Diversity of haplotypes sampled, including the percentage of common and rare haplotypes.

Primer set	**Region**	**Clones Sampled**	**# Distinct Haplotypes**	**% Common haplotypes[a]**	**% Rare haplotypes[b]**
A	150+, 568-	731	57	7.0%	54%
B	1199+, 1765-	653	66	1.4%	78%

[a] Common haplotypes defined as sequenced ≥50 times in the six samples.
[b] Rare haplotypes defined as sequenced ≤2 times in the six samples.

Table 1.3 Haplotypes sampled with 100% identity to published sequences, including number of samples containing the haplotype, and number of clones sequenced with the haplotype

Primer Set	Taxon hit with BLAST search	Genbank Accession Number	Haplotype Number	# Collections Containing Haplotype	Clones Sequenced with Haplotype
A	*Laboea strobila*	AF399153	hap27	3	65
	Strombidium biarmatum	AY541684	hap36	4	184
	Strombidinopsis jeokjo	AJ628250	hap44	3	17
	Pelagostrobilidium neptuni	AY541683	hap40	1	4
B	*Pelagostrobilidium neptuni*	AY541683	hbp20	4	23
	Strombidium biarmatum	AY541684	hbp23	6	251
	Tintinnopsis tubulosoides	AF399111	hbp65	2	8
	Uncultured eukaryote	AY129053	hbp71	1	11

Table 1.4 Chao1 estimates of sample diversity (with 95% confidence intervals) for each primer set

Sample	Chao1 Primer A	Chao1 Primer B
BH Oct	171 (90.9-223.8)	171 (90.9-223.8)
BH May	91 (50.8-196.1)	78 (43.9-130.8)
SI Oct	6 (2.6-20.7)	136 (73.8-191.3)
SI May	325 (161.0-347.7)	253 (129.1-292.8)
CT Oct	105 (58.0-160.3)	153 (82.1-207.4)
CT May	28 (15.8-64.4)	78 (43.9-130.8)

Table 1.5 Ciliate morphospecies identified by light microscopy from May 2005.

Taxon	**BH**		**SI**		**CT**	
	m	**h**	**m**	**h**	**m**	**h**
Choreotrichs						
Tintinnids (*Tintinnopsis* spp., *Eutintinnus* sp.)	3	3	6	7	4	0
Strobilidium, Strombidinopsis, Lohmaniella spp.	3	3	4	3	4-5	0
Oligotrichs						
Strombidium spp.	5	9	4	17	4-5	12
Laboea strobila	1	1	1	1	1	0

m Morphospecies identified by microscopy
h Haplotypes sequenced

Figure 1.1 Map of all collection sites from October 2004 and May 2005.
Symbols representing each site on this map are used in figures 1.2, 1.3, 1.7, and 1.8 to identify location.

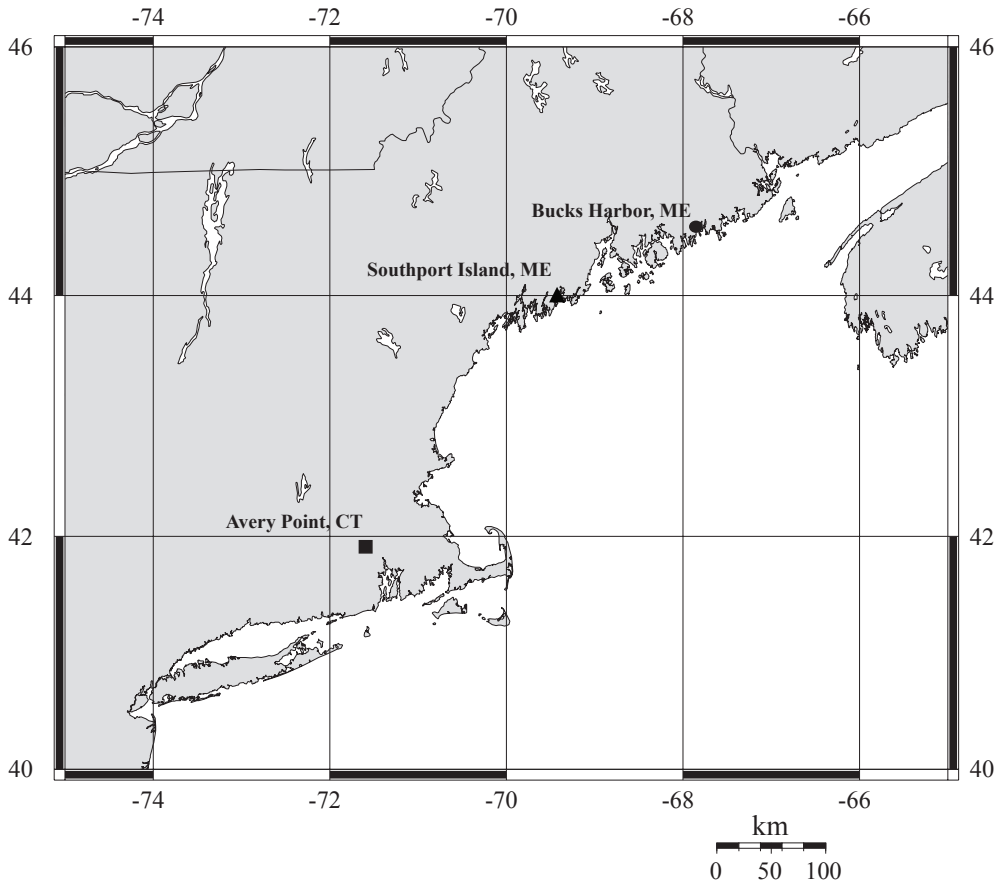


Figure 1.2 Bayesian phylogeny based on the analysis of partial SSU rDNA gene sequences for primer set A.

Topologies shown are trees with the highest likelihood scores. Numbers at nodes are Bayesian posterior probabilities. All branches are drawn to scale. Phylogeny (likelihood score -4546.744) of 57 environmental haplotypes sequenced with primer set A with 34 published sequences, based on 375 characters.

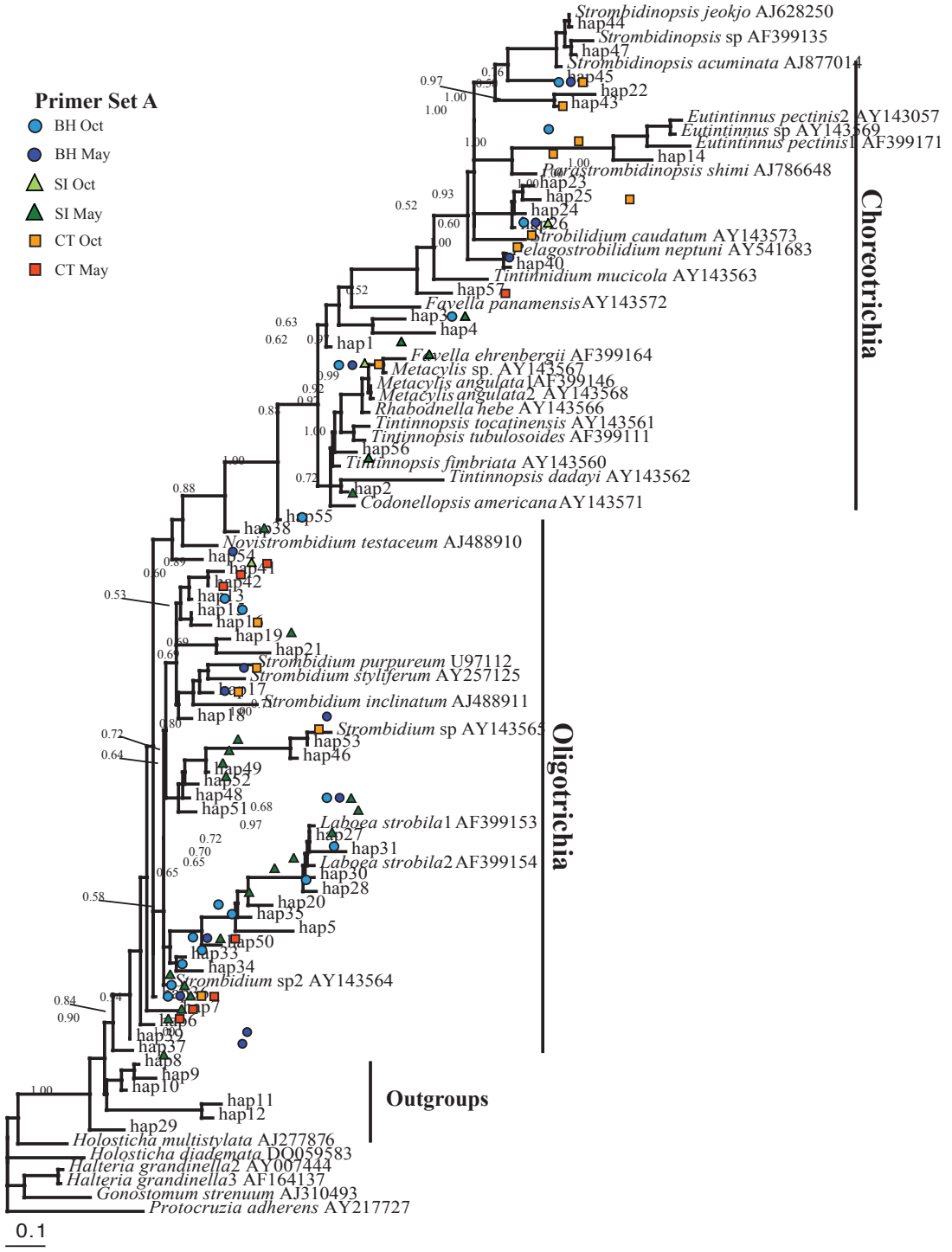


Figure 1.3 Bayesian phylogeny based on the analysis of partial SSU rDNA gene sequences for primer set B.

Phylogeny (likelihood score -4559.868) of 67 environmental haplotypes sequenced with primer set B with 34 published sequences, based on 477 characters. Symbols next to haplotype names indicate where each haplotype was found.

Figure 1.4 Total diversity of ciliate SSU rDNA haplotypes sampled with primer sets A and B using different percent similarity cutoffs.

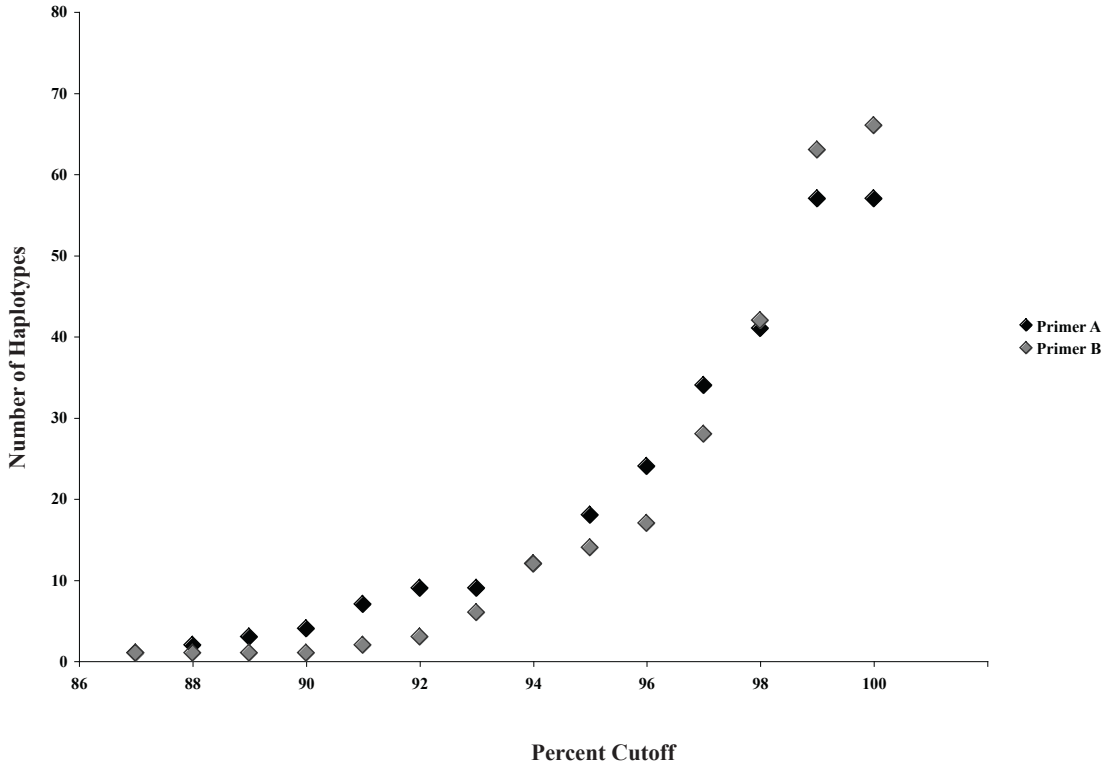


Figure 1.5 Distribution of rare and abundant ciliate SSU rDNA haplotypes across sampling locations and times for primer set A.

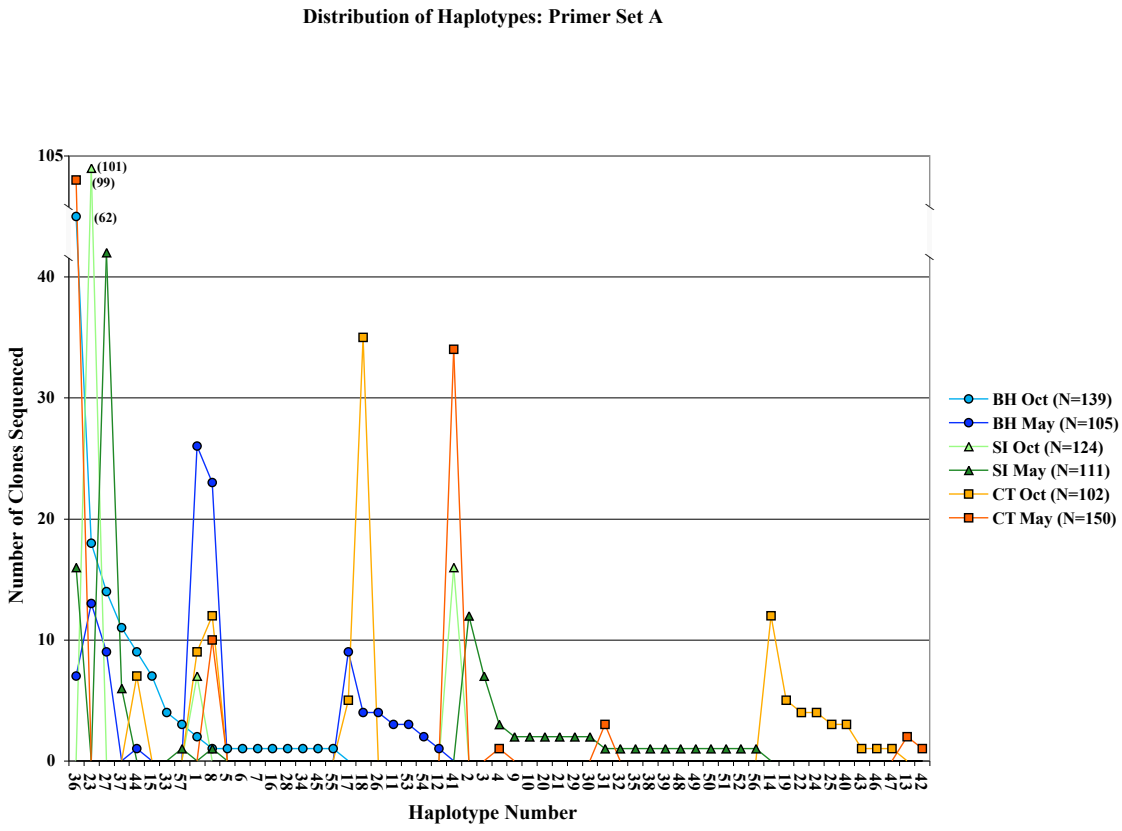


Figure 1.6 Distribution of rare and abundant ciliate SSU rDNA haplotypes across sampling locations and times for primer set B.

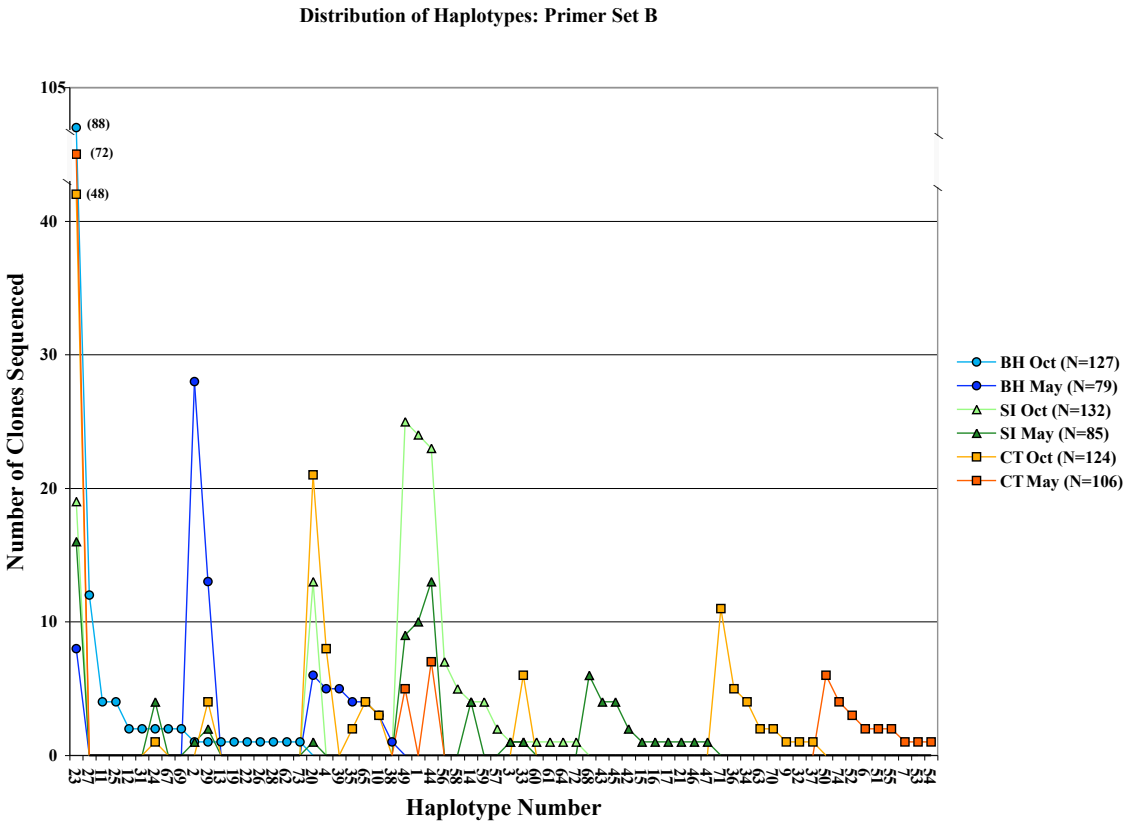


Figure 1.7 Rarefaction curve calculated with EstimateS version 7.5 for primer set A.

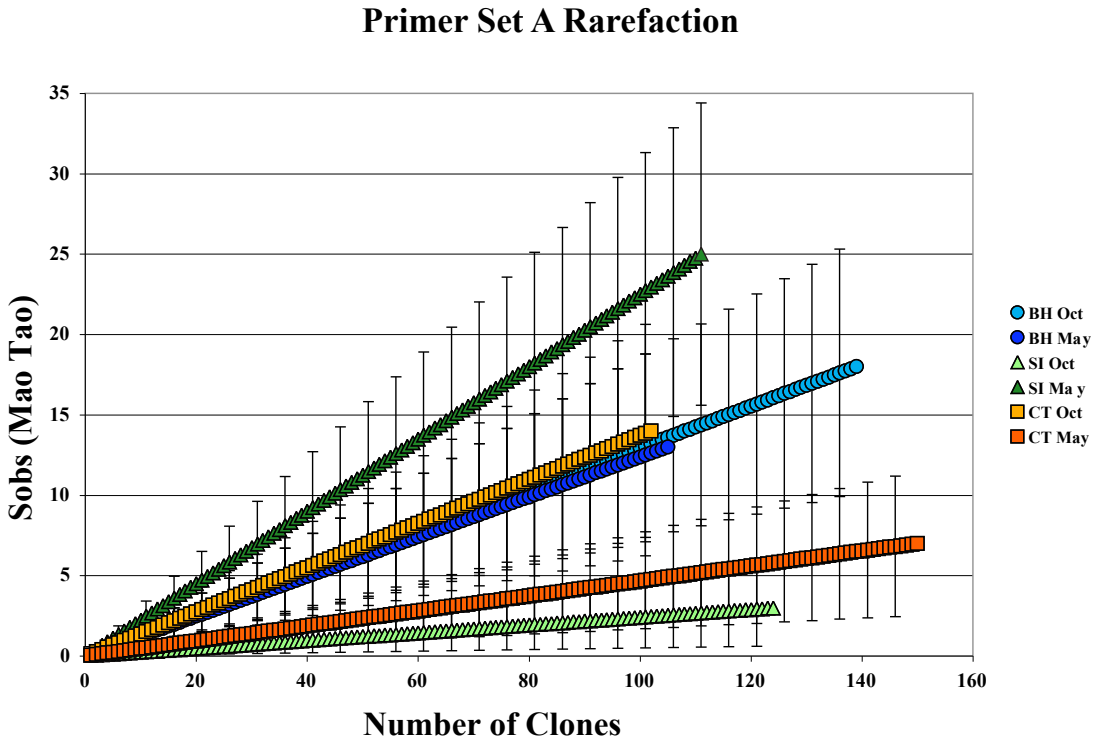
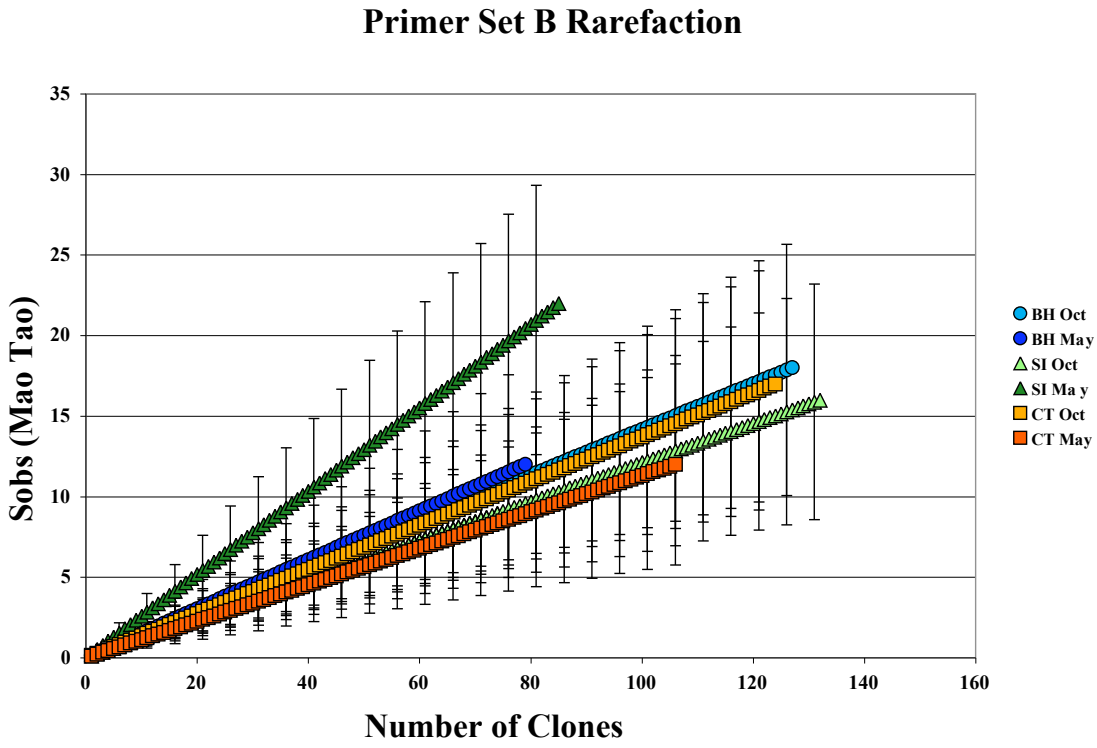


Figure 1.8 Rarefaction curve calculated with EstimateS version 7.5 for primer set B.



CHAPTER 2

A PRELIMINARY STUDY OF THE DIVERSITY OF OLIGOTRICH AND CHOREOTRICH CILIATES IN COASTAL MARINE SEDIMENTS: RELATIONSHIPS WITH OVERLYING PLANKTON

Introduction

One key to understanding the diversity and ecology of two abundant groups of marine ciliates, the Choreotrichia and Oligotrichia, is the relationship between benthic and planktonic forms. While the ciliates in these two groups are predominantly swimmers (Shimeta and Sisson 1999), there is crossover between benthic and pelagic environments for many species. Some taxa are described as epibenthic, living in the layer of water just above the sediment (Fenchel and Jonsson 1988, Shimeta and Sisson 1999), some have the capacity to live attached to sediment particles for a period and then become free-swimming (Jonsson *et al.* 2004), and a large number of taxa within these two groups spend a portion of their life cycle in dormancy, persisting in the sediments in cyst form (Kamiyama and Aizawa 1992, Kamiyama 1994, Kim and Taniguchi 1995, Montagnes *et al.* 2002a, Montagnes *et al.* 2002b, Muller 2000, Muller 2002, Muller 2007, Paranjape 1980, Reid and John 1978, and Reid 1987). An accurate assessment of ciliate dynamics in the plankton requires careful study of both benthic and pelagic environments, and the extent of coupling between the two environments.

The role of the cyst in the life cycle of marine planktonic ciliates is particularly critical for understanding their distribution, evolutionary history, and ecology (Corliss and Esser 1974) as cysts provide a mechanism for dormancy during periods of poor environmental conditions. Relatively few marine ciliate species have been directly

studied to determine conditions for encystment and excystment, period of dormancy (Paranjape 1980, Kamiyama and Aizawa 1992, Kamiyama 1994, Kim and Taniguchi 1995, Kim and Taniguchi 1997), and role of the encystment cycle in the ecology of the organism (Montagnes *et al.* 2002b). Moreover, studies on the conditions related to encystment and excystment in ciliates reveal different patterns and potential causes depending on the species (Paranjape 1980, Kamiyama and Aizawa 1992, Kamiyama 1994, Kim and Taniguchi 1995, Kim and Taniguchi 1997, and Montagnes *et al.* 2002b. While some data link the cycle of encystment with environmental factors such as light (Kamiyama and Aizawa 1992), temperature (Kamiyama and Aizawa 1992, Kim and Taniguchi 1995, and Kim and Taniguchi 1997), and presence of food (Kamiyama 1994), other data suggest a temporal/seasonal cycling independent of external environmental conditions (Paranjape 1980, Kim and Taniguchi 1997, and Montagnes *et al.* 2002b).

A further factor limiting my understanding the role of cysts in the life cycle of ciliates is identification based on the limited morphological features of the cysts, which are highly convergent (Belmonte *et al.* 1997, Foissner *et al.* 2007). In the case of ciliates that encyst within a lorica, as in the tintinnids, this is less of a problem (Reid and John 1978), but for aloricate species, identification is not certain without direct observation of excystment (Reid 1987, Muller 2007). Morphological surveys of ciliates in the benthic environments frequently capture members of the Oligotrichia and Choreotrichia (Shimeta and Sisson 1999, Shimeta *et al.* 2002, Hamels *et al.* 2005, Madoni 2006, Shimeta *et al.* 2007), but are frequently limited to identification at the genus level using morphological approaches.

More is known about planktonic ciliates where morphology provides a wealth of data (Dolan *et al.* 2007) and where molecular studies have revealed tremendous diversity with many rare haplotypes (Doherty *et al.* 2007). Planktonic ciliates show high molecular diversity at the SSU rDNA locus (Katz *et al.* 2005, Doherty *et al.* 2007), and primer sequences have been developed to detect ciliates from environmental samples within the subclasses Choreotrichia and Oligotrichia (Doherty, *et al.* 2007). Ciliates from these subclasses sampled across three coastal locations comprised distinct assemblages with a few ubiquitous and abundant haplotypes (Doherty *et al.* 2007), and many singletons (haplotypes unique to a particular sample).

In this pilot study, I set the groundwork for an alternative to morphological methods for studying benthic assemblages of oligotrichs and choreotrichs and comparing them to those of the overlying water. My goal was to compare genetic diversity between sediment and plankton samples collected from the same locations as a means of assessing the potential of methods for monitoring exchange between these two communities. Specifically, I focus on two major areas. First, I investigated the robustness of molecular approaches to repeated sampling in each environment, plankton and sediment. This determination is critical, particularly for comparing across communities. In each case, I re-sampled ciliate communities to determine the reproducibility of my collection methods, reflecting the degree of spatial heterogeneity in the environment, as well as the reproducibility of the molecular approaches I used in capturing diversity. Second, I investigated the relationship between genetic diversity of ciliate communities sampled in marine sediments and ciliate communities sampled in the plankton.

One hypothesis I test is that the ciliate community observed in the plankton represents a subset of the diversity found in the benthic community, including cysts, beneath it. In this scenario, the plankton community at a given time is based on prevailing environmental conditions, predation, and chance, while the benthic community, including encysted planktonic forms, represents the longer-term diversity in a given region. I tested these predictions on sediment samples collected in the Gulf of Maine and Long Island Sound in May 2005, and compared my results to previously published data from plankton samples collected at the same times and locations (Doherty *et al.* 2007).

Materials and Methods

Sample Collection

Sediment: Ciliates were sampled from three near-shore locations. Two were sampled in the Gulf of Maine: Bucks Harbor, ME (44˚38.20'N, 67˚22.29'W) and Southport Island, ME (43˚49.05'N, 69˚39.16'W), and one was sampled in Connecticut on Long Island Sound (41˚19.00'N, 72˚03.65'W). Sediment samples were collected on the same day in May 2005 in tandem with plankton collections (Doherty *et al.* 2007). In each location, I collected the uppermost ~1cm layer of sediment in a 50ml conical tube (approximately 20-25g total). The Maine samples were immediately placed in a cooler with dry ice during transport to the lab. Connecticut samples were collected at UConn's Avery Point marine campus and did not need to be transported. Samples from both locations were stored at -80˚C until DNA extraction. I recorded ambient water temperature, salinity, and chlorophyll concentration in the water column at each

collection. For chlorophyll, 100 ml was filtered onto a Whatman GFF glass fiber filter. The filter was folded in half, wrapped in aluminum foil and stored at -80°C prior to extraction in 90% aqueous acetone and quantification by fluorescence.

Plankton: Plankton samples for sub-sampling robustness estimates were collected at Southport Island, Maine and Ipswich, MA (42º42.708' N, 70º47.79' W) using a pre-concentration step (siphoning 50-60L of water through a submerged 20μm mesh) as described in Doherty *et al.* (2007). A second sample from Ipswich, MA was collected using the following approach: 2L of water was sampled and filtered through a 3.0μm cellulose nitrate filter (Millipore cat. # 7193-002), immediately placed on dry ice for transport back to the lab, and stored at -80°C until DNA extraction.

DNA Extraction, Amplification, and Sequencing

For each sediment DNA extraction, I weighed ~1g of sediment, and extracted using the DNeasy plant kit by Qiagen (cat. # 69104). I modified the manufacturer's protocol for my sediment by initially placing the sample in either DNA prep buffer (100mM NaCl, Tris-EDTA at pH 8, and 0.5%SDS) or buffer AP1 from the DNeasy kit, mixing by vortexing, and removing the supernatant to use for genomic DNA extraction. Plankton samples were extracted using methods previously described (Doherty *et al.* 2007).

I amplified DNA fragments using polymerase chain reaction (PCR) with Phusion polymerase and reagents from Finnzyme Inc using primers designed to be specific for Choreotrichia and Oligotrichia small subunit ribosomal DNA (SSU) as described in Doherty *et al.* (2007). PCR products were gel-isolated and cleaned using the UltraClean GelSpin DNA purification kit from Mo Bio Laboratories (cat. #12400-100). I used either

the pSTBlue-1 Perfectly Blunt Cloning Kit from Novagen (cat. # 70191-4DFRZ) or the Zero Blunt TOPO PCR Cloning Kit (Invitrogen cat. #45-0245) for cloning, and then picked and miniprepped colonies using the PureLink 96 Plasmid Purification System from Invitrogen (cat. # 12263-018). Sequencing reactions were performed using Big Dye Termination Kit (Applied Biosystems), cleaned with a sephadex plate column, and sequenced on an ABI 377 automated sequencer.

Sequence Assembly and Phylogenetic Analysis

I assembled and edited sequences using SeqMan (DNAStar Inc.). I selected a 99% similarity cutoff for genealogical analyses and diversity estimation to allow for discrimination between highly related but distinct taxa (Doherty *et al.* 2007). Haplotypes were then checked for identity with published sequences using BLAST search on NCBI (www.ncbi.nlm.nih.gov). All sequences in the analysis were screened for PCR chimeras using the recombination detection software RDP version 2.0 (Martin *et al.* 2005) with the Chimaera (Maynard Smith 1992, Posada and Crandall 2001, and Posada 2002) and GENECONV (Padidam *et al.* 1999) applications. Putative recombinants were then visually inspected in MacClade version 4.06 (Maddison and Maddison 1992) for confirmation.

For genealogical analyses, haplotypes were aligned with published sequences from identified morphospecies obtained by searching GenBank for all entries recorded as Choreotrichia and Oligotrichia. In addition, I included 150 sequences from uncultured environmental samples in my phylogeny that appeared in BLAST search results as closely related to known Choreotrichia and Oligotrichia sequences. I used the CLUSTAL W algorithm as implemented in MegAlign (DNAstar Inc., Madison, WI) to

align my sequences with the published sequences. I finalized alignments by eye in MacClade version 4.06 (Maddison and Maddison 1992).

Bayesian analyses were conducted for each primer dataset using a GTR + G + I model of sequence evolution in MrBayes (Ronquest and Huelsenbeck 2003). Four simultaneous MCMCMC chains were run for 10,000,000 generations sampling every 100 generations. Stationarity in likelihood scores (L) was determined by plotting the -1n (L) generation number. All trees below the observed stationarity level were discarded, resulting in a "burn in" of 75,000 generations. Estimation of best-fit models for partial SSU rDNA gene sequences was performed using MrModelTest 2.2 (Nylander 2004).

Statistical Analyses

I estimated haplotype richness in samples by calculating rarefaction curves using EstimateS version 8.0 (Colwell 2006), comparing number of clones sequenced to number of observed haplotypes based on my 99% assembling criterion. I also calculated the non-parametric richness estimator, Chao1, with EstimateS using 100 randomizations, sampling without replacement.

To test whether sediment samples cluster together in the phylogenetic tree based on environment, I calculated principal coordinates analysis (PCoA) and hierarchical clustering analysis using the online software UniFrac (Lozupone *et al.* 2006). UniFrac can be used to determine whether environments differ significantly in community composition, if community differences are concentrated within particular lineages of the phylogenetic tree, or if environmental factors group communities together (Lozupone *et al.* 2006). I used the Bayesian tree and a text file with sequence labels mapped to environmental samples as input for the UniFrac analyses. The distances were plotted as

points in a multidimensional space, one dimension fewer than the number of samples, so that the principal coordinates describe how much of the variation each of the axes in this new space explains. These coordinates were then analyzed for correlation with environmental parameters of the samples. I used the unweighted pair group method with arithmetic mean (UPGMA) hierarchical clustering algorithm, which clusters pairs of samples, and tested robustness of these clusters with jackknife analysis, a non-parametric estimator based on 100 randomized sub-samples. I tested whether the sediment samples differed significantly from one another on the Bayesian tree by conducting a P-test in UniFrac, which estimates similarity between communities as the smallest number of changes that would be required to explain the distribution of sequences in the tree (Martin 2002).

Subsampling

Sediment: To determine reproducibility of sediment diversity estimates, I took three or more 1g subsamples from each collection, and compared the haplotype richness among subsamples. Genomic DNA from each subsample was amplified by PCR in 2 or more separate reactions. These PCR products were cloned and sequenced individually, and resulting haplotype diversity was evaluated for reproducibility by PCR reaction within and between subsamples.

Plankton: To evaluate the robustness of my collection methods in the plankton, I sampled water from Ipswich, MA (42º42.708' N, 70º47.79' W), and compared two filtering approaches. In the first approach, I used the pre-concentration method described in Doherty *et al.* (2007), where a large volume of sea water (60l) was pre-concentrated down to 5l by siphoning through a submerged 20µm mesh, and then filtered. In the

second approach, I sampled a much smaller volume (2l), and filtered it without pre-concentration approximately 1 hour after sampling.

To test the reproducibility of molecular methods in the plankton, I re-sampled genomic DNA from filters collected in the Gulf of Maine in May 2005, and compared the resulting diversity to my previously published estimates from the same samples (Doherty *et al.* 2007). DNA was extracted from the filters in the manner described in Doherty *et al.* (2007), amplified by PCR, and clone libraries were generated. From these clone libraries, 257 clones were sequenced. The results from this additional sequencing effort were compared to initial estimates of diversity obtained for the sample (Doherty *et al.* 2007).

Results

Phylogenetic Diversity in Sediments

In total, I analyzed 729 clones from sediment samples (Table 2.1). I identified 49 haplotypes, and of these 49, more than half (27) were rare in the sample (represented by 3 or fewer sequences). The remaining 22 haplotypes were represented by a greater number of sequences, and all but two were sampled in multiple PCR reactions (Table 2.1). The most abundant haplotype, sampled by 206 clones, was identified through BLAST searches to be 100% identical to an environmental spirotrichid haplotype sampled in New England Coastal waters (GB #EF553401). This haplotype falls within the Choreotrichia on my phylogeny as sister group to a sequenced morphospecies, the tintinnid *Codonella sp.* (GB #. DQ487193; Fig. 2.1). A second haplotype, found in high abundance (126 clones) as well as throughout the samples, was a haplotype that BLAST search results

show to be 100% identical to morphospecies *Strombidium biarmatum* (GB#AY541684) within the Oligotrichia (Fig. 2.2). This morphospecies was also the one most commonly found with molecular methods in planktonic ciliate samples (Doherty *et al.* 2007). It was only recently described taxonomically from the Gulf of Trieste in the Mediterranean Sea (Agatha *et al.* 2005).

Twenty-eight of the 49 haplotypes sequenced from the sediment had been seen in previously published planktonic samples (Doherty *et al.* 2007). Only one of these can be associated with a described morphospecies, the aforementioned *Strombidium biarmatum*. Sixteen haplotypes were found in more than one sediment sample while 33 haplotypes were captured in only one sample ('Singleton haplotypes'; Table 2.1). Of these 33 singleton haplotypes, 20 had been previously captured in plankton samples (Doherty *et al.* 2007), leaving 13 of 49 haplotypes that were found only once among the pooled plankton and benthic observations.

I detected no evidence of PCR recombination in my haplotype sequences. Using the RDP software (Martin *et al.* 2005), no recombinants were detected using the Chimaera program (Maynard Smith 1992, Posada and Crandall 2001, and Posada 2002), even after decreasing the stringency of the test by incrementally raising the p-values. I also applied the GENECONV program (Padidam *et al.* 1999), which applies a sliding window approach to identification of recombinants for every possible triplet of bases. This program did identify putative recombinant sequences, but I determined by visual inspection in MacClade that they that contained levels of polymorphism too high to be consistent with PCR recombination.

Bayesian analyses of the SSU rDNA data from of my sediment samples combined with published data show that the majority of haplotypes in my sediment samples fall within the Oligotrichia (30 of the 49 sequences), 18 haplotypes fall within the Choreotrichia, and one haplotype (hbp110) groups most closely with the outgroup, the Protocruziid spirotrich *Protocruzia adherens* (Figs. 2.1 and 2.2).

Subsampling Reproducibility

Sediment: I tested my collection approach by examining replicate subsamples from the same initial collection of sediment (~20-25g; Table 2.1). Levels of diversity and haplotype representation varied widely among these replicates (Tables 2.1 and 2.4). For example, comparing replicates with ~20 clones sequenced, subsample #2 showed a diversity of 1-2 haplotypes, while replicate #3 revealed a diversity of 6-7 different haplotypes (Table 2.1). Chao1 and rarefaction diversity estimates calculated for the samples also varied between replicates (Table 2.4; Fig. 2.3). For the sake of clarity, I show the rarefaction curves estimated for only one of the locations, Southport Island, ME, to illustrate the inconsistency between subsamples (Fig 2.3).

To address whether the variance observed was due to PCR bias, I analyzed replicate PCR reactions on DNA extracted from a single subsample. My diversity estimates showed more consistency in replicate PCR experiments conducted on the same DNA extraction, than on replicate extractions performed on sediments in the same location (Table 2.4; Fig. 2.3). Comparisons of membership between these replicate subsamples are consistent with estimated diversity results.

Plankton: In contrast to sediments, subsamples of haplotype diversity from plankton communities show greater similarity between subsamples (Tables 2.2 and 2.3).

Because of my concern that sampling methods might bias diversity estimates, I compared plankton samples collected using different filtering methods from the location in Ipswich, MA (Table 2.3). These samples, Standard Collection (C_{std}) and Modified Collection (C_{nov}), are similar in that they are both dominated by the same abundant haploypte, which I call hbp95, and they share 50% of their haplotype assemblages. The difference between these samples is largely due to presence/absence of rare haplotypes. One notable exception is haplotype 258_05, which was relatively abundant in the 60L pre-concentrated sample, but rarer in the 2L sample, suggesting that this haplopyte may have died off rapidly in the two hours between collection and filtering.

I further assessed the impact of greater depth of sampling on one of my planktonic communities by analyzing an additional 257 clones from Southport Island, ME (M_4) and compared these to my published data from the same location (M_1) (Table 2.3). Again, my estimates of membership and diversity in this planktonic sample appear robust as similar estimates were produced by these different levels of effort. Rarefaction curves generated from the initial 84 sequences and the additional 257 sequences sit directly on top of one another, indicating identical estimates of diversity between the samples (Fig. 2.4).

Using Fisher's exact test, I investigated whether the samples amplified in repeated molecular approaches and the samples collected using different water sampling were drawn from the same distribution. The results from this test strongly support the null that the samples are independent of one another ($p < 0.0001$), indicating that the observed overlap in plankton subsamples is not significant. I suspect that the large proportion of rare haplotypes in these datasets contributes to these differences.

Comparison of Genetic Diversity in Sediment and Planktonic Samples

To determine the relationship between sampling environment and ciliate assemblages, I compared phylogenetic distances for each sample with those sampled in other locations. Using the software in Unifrac (Lozupone *et al.* 2006) I generated an environmental matrix using genetic distances in the Bayesian SSU rDNA tree associated with sampling location. I performed the analysis using both weighted and unweighted branch lengths to determine the effect of abundance on the clustering of haplotypes, versus presence/absence. I discerned a pattern only in the case where I used unweighted branches, which is a qualitative (presence versus absence) rather than a quantitative assessment. Principal coordinates analyses using unweighted branches group the sediment communities together, distinct from plankton communities collected in the same locations at the same time (Fig. 2.5). Hierarchical clustering using UPGMA is consistent with these findings, but jackknife analyses show moderate to weak support for many of the nodes (Fig. 2.6). Moreover, analyses using weighted branch lengths cause the observed clustering pattern to fall apart. Hence there is a weak relationship between sediment communities based on membership, but not on numerical dominance or rarity.

Plankton samples do not cluster with their respective benthic samples. There were low levels of overlap between sediment and plankton assemblages (Table 2.5) (Doherty *et al.* 2007). While the total number of sediment haplotypes captured at each location ranged between 17-32, and the plankton haplotypes range between 24-47, the maximum overlap between plankton and sediment at any given location was only 3-5 haplotypes (Table 2.5). A much higher level of overlap of haplotypes was found among spatially separated samples for both plankton and sediments (24 and 15 overlapping haplotypes, respectively) than between plankton and sediment at the same location.

Discussion

Plankton diversity estimates are robust to varying collection methods and to subsampling (Tables 2.2 and 2.3). My standard sampling practice, which involves immediately filtering and preserving a large volume (50-60L) of water after concentration (Doherty *et al.* 2007) gave similar results to a 2L sample processed 2 hours after collection (Fig.2.4; see methods for further details). Similarly, my estimates of diversity and membership are similar based on independent estimates from 84 or 257 clones. The more intensive sampling resulted only in a greater number of rare haplotypes in the sample and the resulting distribution thus differ by the conservative Fisher's exact test (Table 2.2; Fig. 2.4). Together, these results give us greater confidence in my ability to capture the dominant members of ciliate communities in planktonic environments.

In contrast, sediment samples show a high degree of heterogeneity among subsamples in both diversity and membership (Table 2.1, Table 2.3, Fig. 2.3). I don't believe this result to be an artifact of my molecular sampling methods for the following reasons: 1) I find consistent results between different PCR reactions amplifying genomic DNA from the same subsample; 2) the trend is consistent across 26 total PCR reactions for a total of 15 grams of sediment; and 3) the pattern I observe in plankton samples is quite different, suggesting that my molecular methods are robust (see above).

A further consideration in comparing samples taken from the sediment to those collected in the plankton is the difference in spatial scale I are evaluating in each case. I collected sediments by isolating ~20 gram of the top few millimeters of sediments at a single point. The variance in estimates of diversity from individual grams subsampled from the original samples could indicate that Oligotrich and Choreotrich ciliates are very rare in sediment samples. However, the high haploypte numbers in some subsamples

(e.g. 6 haplotypes in subsample BH2a/2b; 9 haplotypes in subsample SI3a/3b and 11 haplotypes in subsample CT2a/2b; Table 2.1) suggest that this is not the case. Instead, my data suggest that there is considerable spatial heterogeneity in ciliates in sediments on the scale of < 1 centimeter. Clearly, additional finer grained sampling strategies are needed to clarify the issue of spatial heterogeneity in ciliates in marine sediments.

Comparison of Genetic Diversity in Sediment and Planktonic Samples

With the important caveat that my sampling of sediments did not produce consistent results between samples, I assessed the similarities in communities between plankton and sediment. I did not observe strong similarities between sediment haplotype assemblages (this study) and the plankton haplotype assemblages in the waters above (Table 2.5) (Doherty *et al.* 2007). Comparisons of relative diversity reveal little overlap between plankton and sediment communities from the same locations (Table 2.5). Using pooled sediment subsamples as a proxy, I find comparable levels of diversity in sediments as compared to plankton samples (Table 2.4), and there is no evidence that sediments are sources of broader genetic diversity from which the plankton community is drawn, as a 'seed bank' hypothesis for benthic assemblages would predict. Also inconsistent with a seed bank hypothesis, the sediment communities are more similar to each other in clustering analyses than they are to the community in the plankton directly above, though support here is weak (Figure 2.5). Clearly, the high heterogeneity among subsamples of sediment indicate the need for additional sampling to disentangle these concepts further.

Within sediments, I would expect to find interstitial ciliates, ciliates in cyst form, plus epibenthic ciliates in the small fraction of water taken along with each sample.

Given that very few of the haplotypes I captured are not identical to sequenced morphospecies, I cannot discern between these three possible sources of ciliate diversity in my dataset. Morphological surveys, where identification is often only to the genus level, generally report only 2-4 different oligotrich and choreotrich ciliate types in a sediment sample, although they may be numerically abundant (Fenchel 1969, Patterson *et al.* 1989, Shimeta and Sisson 1999, Shimeta *et al.* 2002, Hamels *et al.* 2005, Madoni 2006, Shimeta *et al.* 2007). My molecular sampling efforts reveal much higher levels of diversity (up to 32 haplotypes at a single site and 49 haplotypes total across three sediment sites) (Table 2.5), indicating either that my efforts are effective at capturing a good portion of cysts in the sediment, or that I are sampling a diversity of cryptic benthic dwelling ciliates.

On the largest scale, I found that common haplotypes were widespread. For example, EF553401, *Strombidium biarmatum*, hbp95, and EF553452 were found at the Connecticut and both Maine sites, a total range of approximately 700 km. However, on the scale of repeated subsampling (~1cm), I found surprising lack of coherence in the presence of different haplotypes. EF553401, for example, represented about half of all sequenced clones from Buck's Harbor subsample 1, yet it was found in none of the other three subsamples at all. This is consistent with the idea that benthic ciliate species are distributed in a very patchy manner on small scales, as indicated by morphologically-based observations (Madoni 2006; Shimeta *et al.* 2007).

While resting stages in other species such as copepods represent a historical record about the genetic makeup of a community (De Stasio 1989, Hairston *et al.* 1996, Marcus *et al.* 1994, and Caudill and Bucklin 2004), I found no evidence that ciliate

resting stages play the same role. Studies of encystment and excysment within the Choreotrichia and Oligotrichia report relatively short periods of dormancy, ranging from 19 hours in *Strombidium occulatum* (Montagnes *et al.* 2002b) to 6 months in *Pelagostrobilidium sp*. (Muller 2002) or two seasons in *Strombidium conicum* (Kim and Taniguchi 1997). The majority of the sediment haplotypes that I sampled, the bulk of which do not match to any known morphospecies, were neither widespread nor abundant in the plankton, with the exception of *Strombidium biarmatum* (Agatha *et al.* 2005), which is a cyst forming species found throughout sediment and plankton samples (Doherty *et al.* 2007).

This preliminary survey found little overlap between benthic ciliate assemblages and those of the overlying water, and no evidence that the benthos serves as a reservoir of diversity for the plankton. I did find similarity in benthic and planktonic assemblages in that both contain a few common haplotypes and many rare ones. This confirms the findings of a number of contemporary studies indicating a much higher degree of diversity in marine eukaryotic microbes than has heretofore been appreciated. Further studies of the degree to which sediment-associated Choreotrichs and Oligotrichs may be interstitial, epibenthic or freely exchanging between sediment and plankton will be needed to uncover the ecological roles of the many haplotypes I observed.

Table 2.1 Haplotypes characterized by PCR of DNA isolated from multiple grams of sediments: Bucks Harbor, ME, Southport Island, ME, and Groton, CT.

Sequence Name	**BH**							**SI**							**CT**											
	1	**2a**	**2b**	**3a**	**3b**	**4a**	**4b**	**1**	**2a**	**2b**	**3a**	**3b**	**4a**	**4b**	**1a**	**1b**	**2a**	**2b**	**3**	**4**	**5a**	**5b**	**6a**	**6b**	**7a**	**7b**
EF553401	67							10			6	6		9							17					
Strombidium biarmatum AY541684	1	6	10	8		11	22	11				2	6	3	1		5	1					9		9	20
hbp95	5							12			1			1	1		45	3		1						
EF553452					20				20	20							1	1								
hbp94	7							24			9		1	3												
hbp110															20	2										
EF553421								3									8	5								
EF553411	3							12																		
hbp93	5							6						4												
EF553415		3	3	6																			2			
hbp92	6							3			1	1		1												
hbp87	4							2			2	1		2												
hbp97	2							4	1					4												
hbp114																	6	2								
hbp96	4										1	1		1												
EF553454		1	4																							
260_06		4	1																							
hbp109	1							3																		
hbp84	2		1																							
hbp112								1											2							
Singleton Haplotypes	23	2		10	13							3		1			15						3		2	
Total Clones Sequenced	130	16	19	24	33	11	22	182	21	20	20	14	7	29	22	2	80	12	2	1	17	0	14	0	11	20

Table 2.2 Comparable haplotypes obtained in SI plankton samples from replicate PCR reactions.

Sequence Name		Initial (M_1)	Resampled(M_4)
Strombidium biarmatum	AY541684	16	64
	EF553425	13	33
	EF553429	9	28
	EF553396	4	23
	EF553391	10	17
	EF553438		14
Pelagostrobilidium neptuni	AY541683	1	11
	EF553430		11
	EF553448	6	5
	EF553401	1	9
	EF553406	4	
	EF553424	4	
	EF553426	4	
	EF553453		3
	EF553399	1	2
	EF553423	2	1
	EF553421		2
	EF553411	2	
	EF553436		2
	EF553439		2
	EF553428	1	
	EF553413		1
	EF553431		1
	EF553412	1	
	EF553450		1
	EF553434		1
	EF553403	1	
	EF553397	1	
	EF553397	1	
	EF553427	1	
	EF553415	1	
Unique Haplotypes			26

Table 2.3 Comparable haplotypes obtained in Ipswich, MA using two collection methods.

Sequence Name	**2L unconcentrated (C_{std})**	**50L preconcentrated (C_{nov})**
hbp95	16	21
hbp114	32	1
Strombidium biarmatum AY541684	8	10
259_04	7	
258_05		5
Pelagostrobilidium neptuni AY541683	1	3
hbp119(Sedsub01)	3	
hbp111	1	2
EF553406		2
AF553452		2
hbp105	2	
EF553421	1	
EF553403	1	
262_14	1	
263_04		1
EF553409		1
262_11	1	
Unique Haplotypes	7	

Table 2.4 Estimates of diversity of ciliates in different sediment samples and from replicate PCRs

	Subsample	Clones	Haplotypes	Chao1 Diversity Estimate (95% CI)
Bucks Harbor, ME				
	1	130*	21	24 (21.54-37.81)
	2a	16	5	5.5 (5.03-13.44)
	2b	19	5	6 (5.07-18.5)
	3a	24	5	6 (5.07-18.5)
	3b	33	4	4 (4-4.06)
	4a	11	1	-
	4b	22	1	-
Southport Island, ME				
	1	182*	12	12.5 (12.03-20.44)
	2a	21	2	2 (2-2.01)
	2b	20	1	-
	3a	20	6	10.5 (6.49-46.92)
	3b	14	7	11 (7.56-35.72)
	4a	7	2	2 (2-2.55)
	4b	29	10	18 (10.99-74.82)
Groton, CT				
	1a	22	3	4 (3.08-15.92)
	1b	2	1	-
	2a	80	11	17 (11.95-48.93)
	2b	12	5	7 (5.18-27.13)
	3	2	1	-
	4	1	1	-
	5a	17	1	-
	5b	0	0	-
	6a	14	4	4.25 (4.01-8.73)
	6b	0	0	-
	7a	11	3	4 (3.08-15.92)
	7b	20	1	-

* 2 or more PCR reactions pooled

Table 2.5 Haplotypes Shared Between Plankton and Sediment

	Plankton		Sediment		Shared
	Clones†	Haplotypes††	Clones	Haplotypes	Haplotypes
Bucks Harbor, ME	213	24	255	32	5
Southport Island, ME	472	47	293	17	5
Groton, CT	229	27	181	19	3
Totals	**914**	**76**	**729**	**49**	**13**

† number of clones sequenced †† number of shared haplotypes in the sample at 99% similarity

Figure 2.1 Phylogeny of Choreotrich haplotypes based on Bayesian analysis of partial SSU rDNA gene sequences.

Topologies shown are trees with the highest likelihood scores. Numbers at nodes are Bayesian posterior probabilities. All branches are drawn to scale (likelihood score -4546.744) based on 477 characters. Locations are indicated by a star shaped symbol, Bucks Harbor, ME in black (★), Southport Island, ME in gray (✯), and Groton, CT in white (☆) to the right of the haplotype. Plankton samples collected in Ipswich, MA are indicated by a triangle symbol (▲).

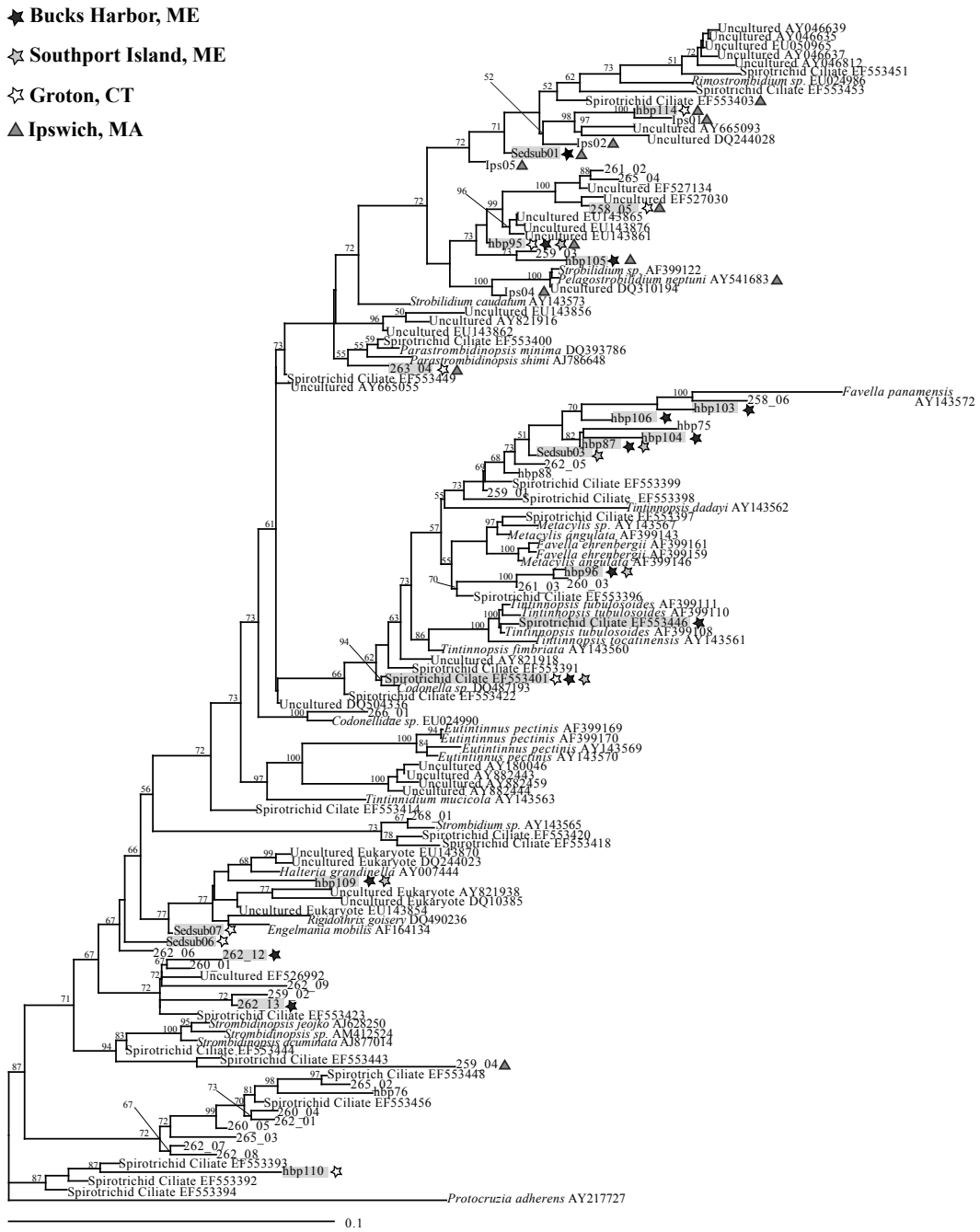


Figure 2.2 Phylogeny of Oligotrich haplotypes based on Bayesian analysis of partial SSU rDNA gene sequences.

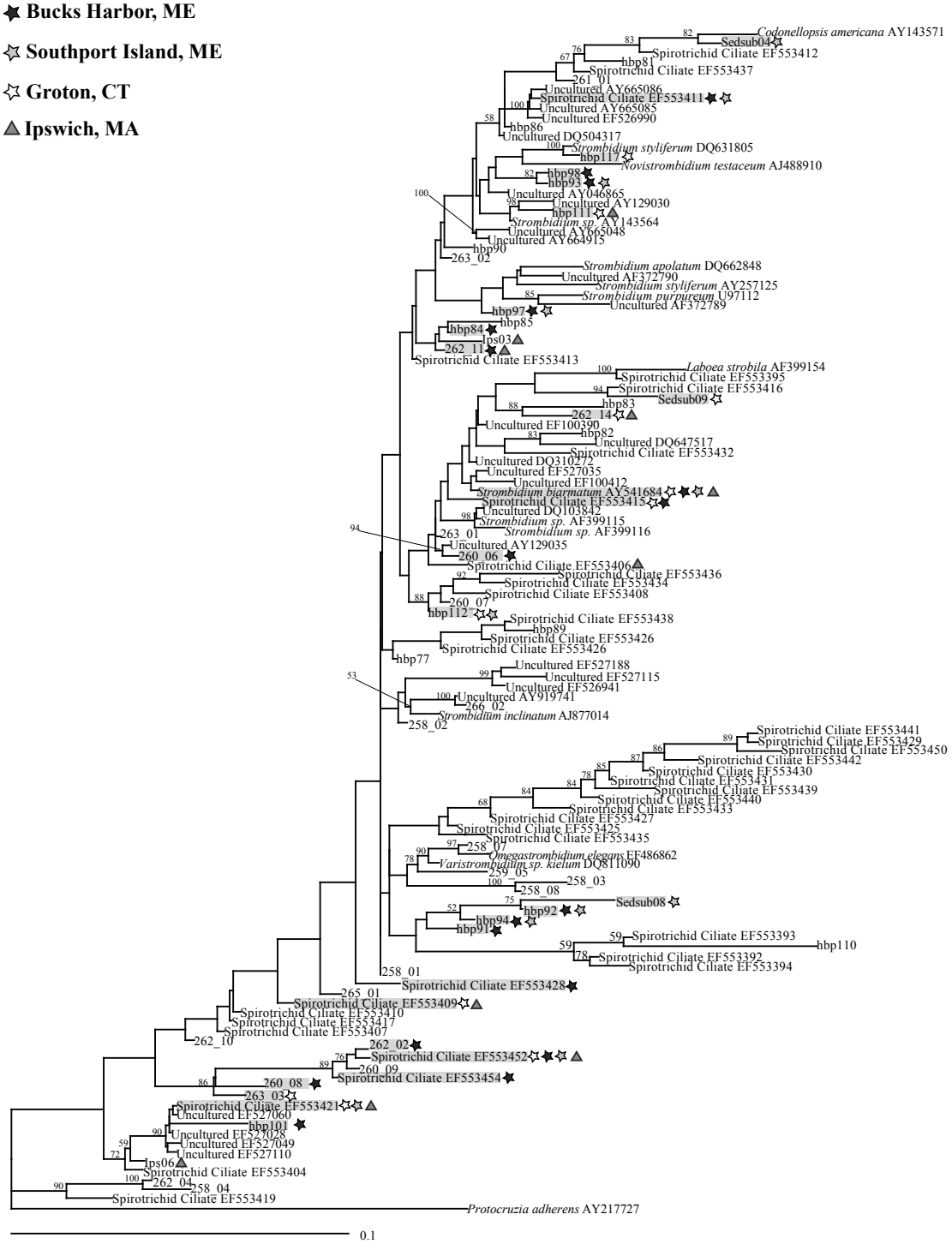


Figure 2.3 Comparisons of diversity as estimated by rarefaction for subsamples calculated with EstimateS version 8.0 for sediments subsampled in Southport Island, ME

Estimated diversity (Sobs, MaoTao) versus the number of clones sequenced is shown. Error bars indicate 95% confidence intervals. Each replicate of sediment sampled is numbered 1-4, while duplicate PCR reactions for the same replicate are indicated by letters (a and b).

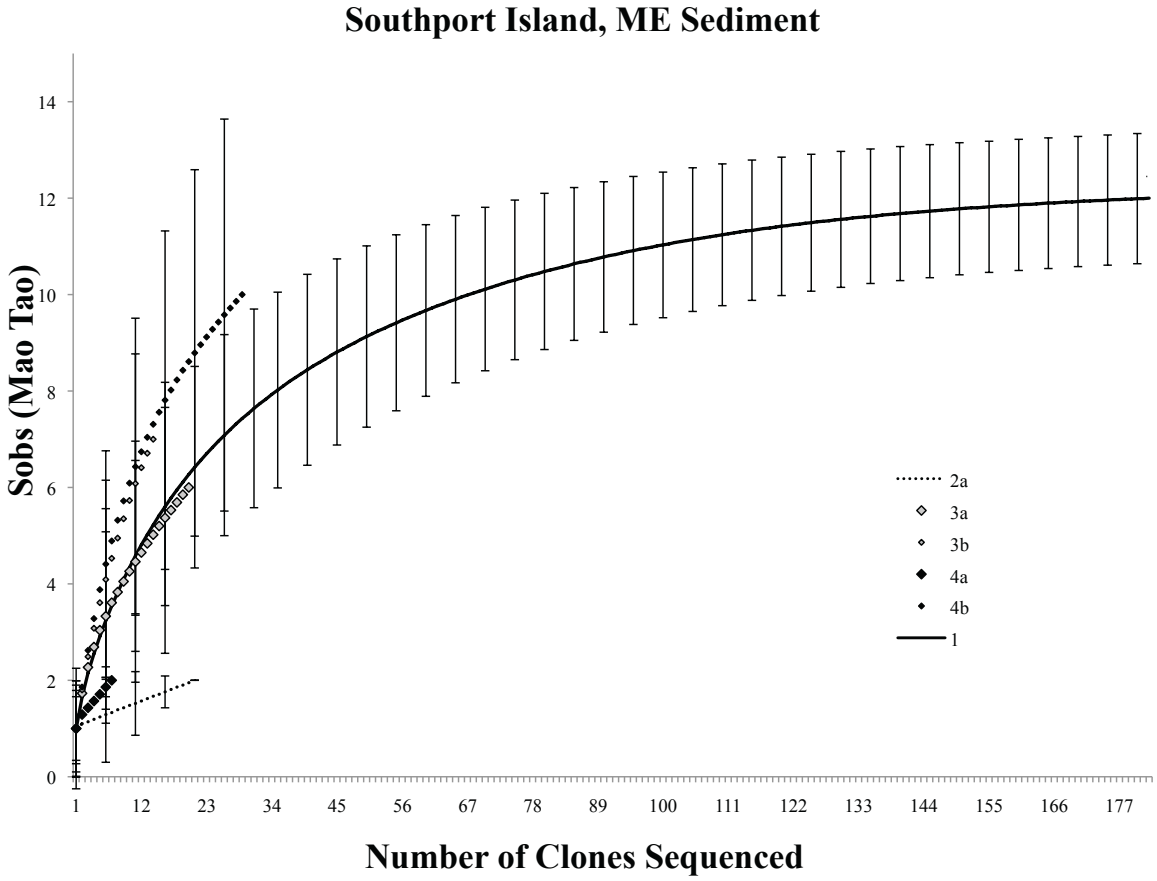


Figure 2.4 Comparisons of diversity as estimated by rarefaction for subsamples calculated with EstimateS version 8.0 for plankton subsamples from Southport Island, ME and Ipswich, MA.
Comparison of diversity between samples taken in Ipswich, MA in 2006, one using a preconcentration step (C_{std}), and one with no preconcentration (C_{nov}). Diversity estimates for Southport Island, ME 2005 samples. Estimated haplotype diversity within an initial clone library of 84 sequences (M_1) compared to a clone library of 257 sequences (M_4) re-sampled from the same DNA.

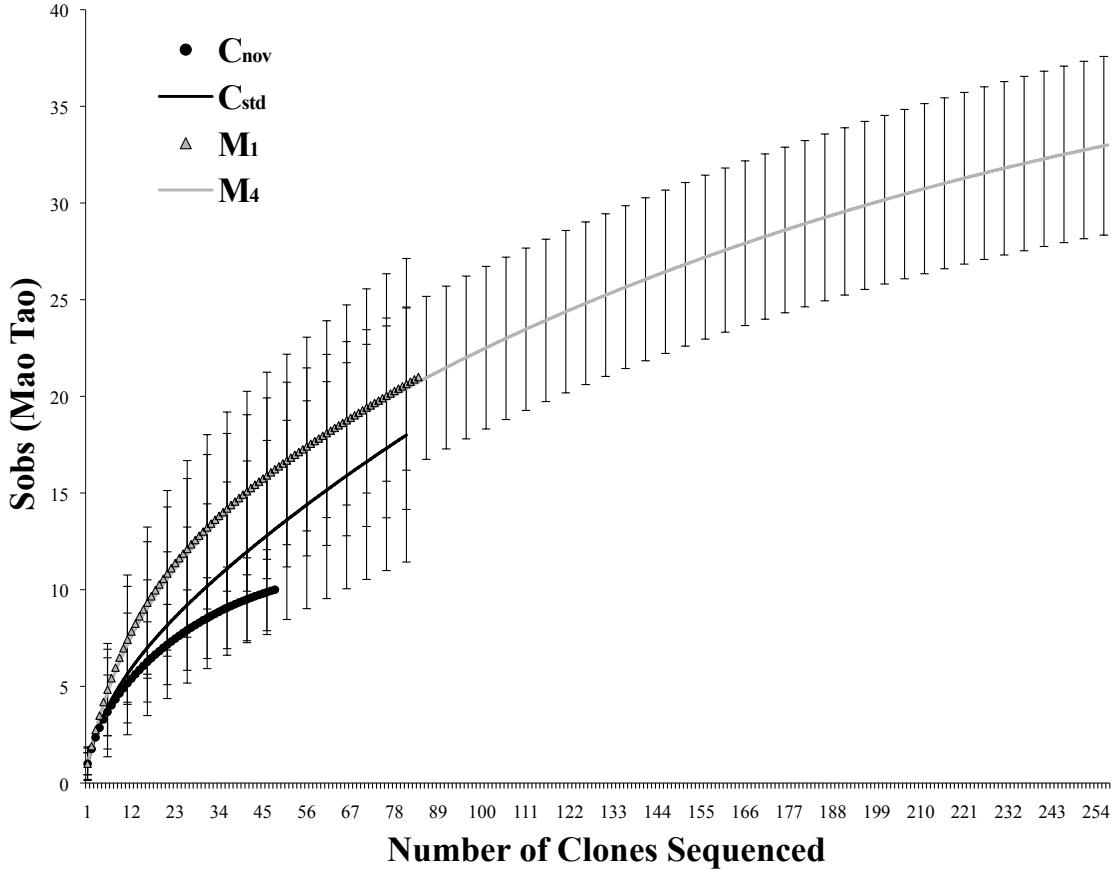


Figure 2.5 Results for Principle Coordinates Analysis (PCoA) based on branch lengths in the Bayesian tree and environmental data for the three sediment locations along with data collected in plankton.

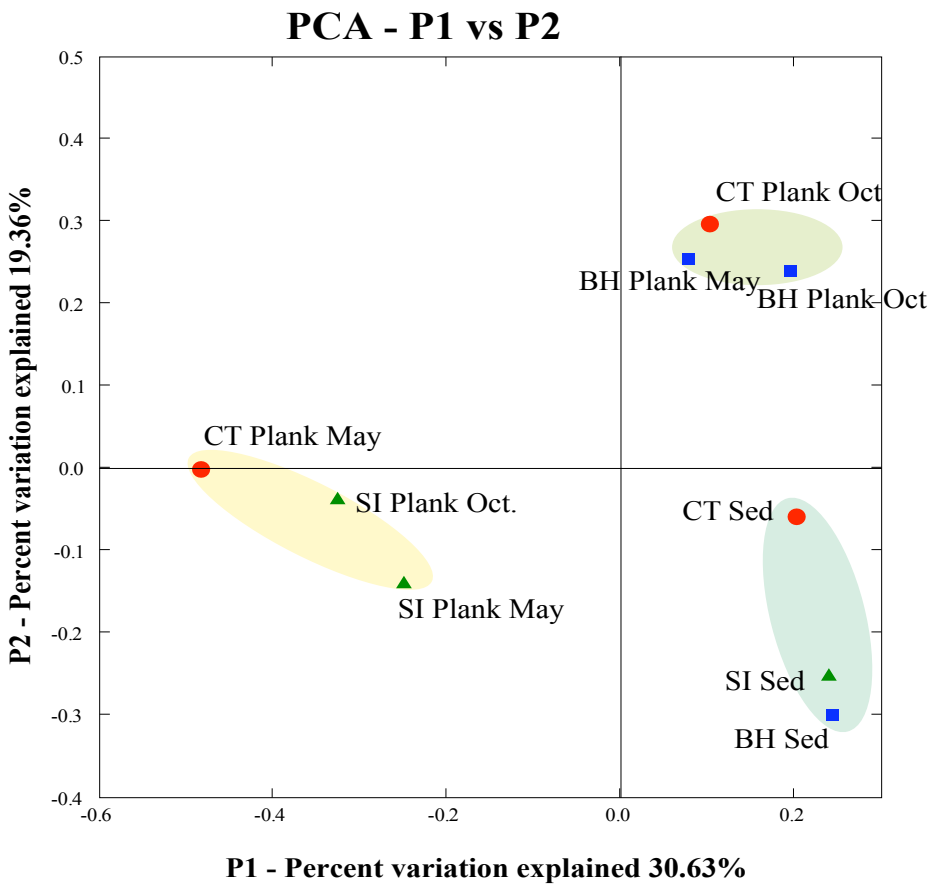
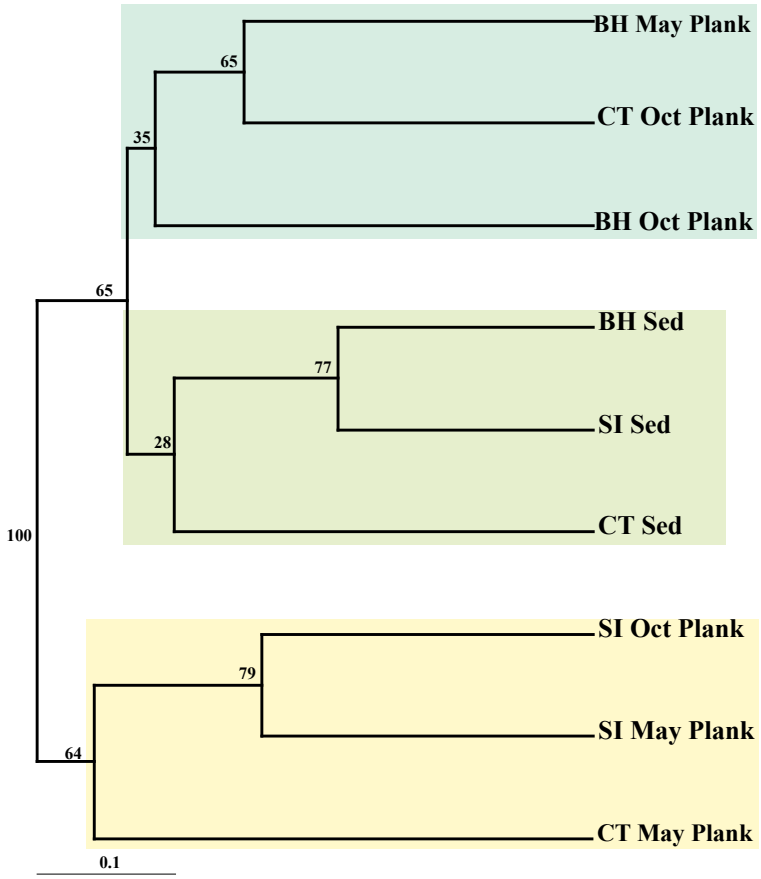


Figure 2.6 Results from hierarchical clustering using UPGMA for the twelve stations. Numbers at the nodes indicate jackknife support from 100 replicates.



CHAPTER 3

CILIATE DIVERSITY AND DISTRIBUTION ACROSS AN ENVIRONMENTAL AND DEPTH GRADIENT IN LONG ISLAND SOUND, USA

Introduction

Using molecular tools, it is now possible to characterize microbial diversity in marine systems that had not previously been measured using morphological methods (Diez *et al.* 2001, Lopez-Garcia *et al.* 2001, Moon-van der Staay *et al.* 2001). More importantly, these tools can be used to test hypotheses regarding distribution patterns of microbes in relation to ecological variables (Countway *et al.* 2007, Not *et al.* 2007, Cuvelier *et al.* 2008). Through testing these hypotheses, I can evaluate how microbial communities are assembled and what processes maintain these assemblages.

An increasing number of studies are using molecular data to test hypotheses regarding microbial distributions, and have reported structuring of microbial communities based on the environment (Fuhrman *et al.* 2006, Countway *et al.* 2007, and Fuhrman *et al.* 2008). For example, research on protistan diversity in marine plankton environments have characterized distinctive shallow and deep-sea assemblages, with different protist communities in euphotic versus bathypelagic zones (Countway *et al.* 2007). Marine bacterioplankton communities studied over a period of 4.5 years were demonstrated to recur in a repeating temporal pattern that is predictable based on a prevailing set of environmental factors (Fuhrman *et al.* 2006). Additionally, marine bacterioplankton reveal a latitudinal gradient of species richness similar to that of larger organisms (Fuhrman *et al.* 2008).

Tools are now available to directly compare phylogenetic diversity of lineages with environmental parameters to determine whether genotypes cluster together by environment (Lozupone *et al.* 2006). Broad-scale analyses using these approaches have uncovered patterns of environmental sorting in a variety of prokaryotic communities (Lozupone and Knight 2007, Lozupone, *et al.* 2007).

Research on ciliate distributions using morphological approaches reveals different abundance distribution patterns in open ocean environments (Dolan *et al.* 2007) and coastal environments (Sitran *et al.* 2009). Species abundance distribution data for tintinnids sampled at 22 stations across the SE Pacific Ocean most closely matched a log series distribution, consistent with the neutral theory of community assembly (Dolan *et al.* 2007). Neutral community assembly models predict random immigration, births and deaths as determining the relative abundance of taxa in a community, in contrast to a niche assembly model where the abundances of different taxa are structured by environment (Hubbell 2001, McGill *et al.* 2007, and Sloan *et al.* 2007). Underlying the concept of neutral theory is the hypothesis of ecological equivalence (Hubbell 2005 and Hubbell 2006), where, for the Dolan *et al.* example, many tintinnid species fill the same ecological roles and stochastic processes such as dispersal determine their distributions in any given sample. In contrast to this pattern, species abundance distributions for tintinnids sampled in a coastal environment in the Western Mediterranean Sea show a log normal pattern, indicating that the community in this sample is shaped by underlying environmental conditions rather than neutral dispersal of species (Sitran *et al.* 2009).

In an earlier study, I designed primers to sample SSU rDNA diversity of ciliates in the Choreotrichia (including the tintinnids) and Oligotrichia, the predominant ciliate

clades in marine plankton (Doherty *et al.* 2007). Molecular data on ciliate communities based on SSU rDNA revealed assemblages of haplotypes differing in composition in ecologically similar environments (Doherty *et al.* 2007). In samples collected hundreds of kilometers apart, a small number of ciliate haplotypes appeared to be both abundant and ubiquitous, but a much higher number of rare haplotypes did not overlap between samples (Doherty *et al.* 2007). These rare types cannot at present be linked to named species and the factors governing their presence or absence is unclear.

For the present study, I chose to examine changes in diversity on a much smaller scale (meters to kilometers) and to evaluate whether changes in ciliate assemblages could be correlated with environmental factors in a coastal area with gradients in salinity caused by river input. I also assessed how the patterns previously observed in tintinnids using morphological approaches relate to patterns revealed using molecular approaches. I sequenced ciliate communities sampled from two depths at each of six stations in Long Island Sound, and determined the SSU haplotype diversity within the sites. My aim was to test the hypothesis that hydrographic properties structure ciliate assemblages over this scale.

Materials and Methods

Collection

In June of 2007, twelve samples were collected at six stations across an increasing depth gradient in Long Island Sound (Fig.3.1, Table 3.1). The locations for sampling were designed to capture a region of Long Island Sound where water exiting the Connecticut River forms a shallow plume of low-salinity water near the surface. The

samples were collected at an increasing distance from the shore, and over increasing depths on either side of a mixing front between fresher and more saline water. The vessel was equipped with an instrument package containing CTD and breakwater sensors and an acoustic doppler current profiler (ADCP). This array was used in two short surveys to better define the presence and dynamics of the plume front. Water samples of 2 l volume were collected at each of the six stations from both surface (0.25-1m), and deeper waters (4m) (Table 3.1). Each sample was reverse-concentrated by siphoning away water through a submerged 20µm mesh and filtered onto a 5µm Millipore cellulose nitrate filter, which was preserved in 1.0ml DNA prep buffer (100mM NaCl, EDTA, and 0.5%SDS) as in Doherty *et al.* (2007). Temperature, salinity, and oxygen concentrations were recorded at each collection using a SeaBird CTD (Table 3.1). Samples were taken at the same time for chlorophyll and accessory pigment analysis to evaluate community composition of the phytoplankton, presumed food of the ciliates. For those samples, 100 ml was collected on glass fiber filters and extracted in acetone. Extracts were analyzed for pigment composition by HPLC according to Van Heukelem et al (1994). To convert the measured pigment concentrations into estimates of the relative contributions of different algal classes to the total phytoplankton biomass, I used the computer program CHEMTAX (version 1.95; Mackey et al 1996; S. Wright, Australian Antarctic Division, pers. comm.). Seed values for the accessory pigment: chlorophyll-a ratios relevant to Long Island Sound phytoplankton were kindly provided by Dr. Judy Li (NOAA/NMFS Milford Laboratory).

DNA Extraction and Amplification

Genomic DNA was extracted from the samples using the DNeasy plant kit by Qiagen (cat. # 69104), with a few modifications. Prior to extraction, I added 0.5µl proteinase K to each filter in 1.0ml of buffer, vortexed, and incubated overnight at 50˚C. The contents were vortexed after incubation, and the filter was removed with sterile forceps. I removed 200µl of the remaining liquid and extracted genomic DNA from this volume according to the protocol in the DNeasy plant kit.

I amplified the targeted region with the polymerase chain reaction (PCR) using methods described in Dohery *et al.* (2007). PCR products were cleaned using the microCLEAN DNA clean-up reagent (Cambio cat. #MZ-1591). I used the Zero Blunt TOPO PCR Cloning Kit (Invitrogen cat. #45-0245) for cloning, and then picked and miniprepped 96 colonies per sample using the PureLink 96 Plasmid Purification System miniprep kit (Invitrogen cat. # 12263-018). Sequencing reactions were performed using the Big Dye Termination Kit (Applied Biosystems), cleaned with a sephadex plate column, and sequenced on an ABI 377 automated sequencer.

Sequence Assembly and Analysis

I assembled and edited sequences using SeqMan (DNA Star). I selected a 99% similarity cutoff for genealogical analyses and diversity estimation to allow for discrimination between highly related but distinct taxa as based on prior studies (Doherty *et al.* 2007). Haplotypes were then checked for identity with published sequences using BLAST search (Altschul *et al.* 1997) on the NCBI website. Sequences were aligned with published sequences obtained by searching GenBank for all entries recorded as Choreotrichia and Oligotrichia. In addition, I obtained 143 closely related environmental

sequences that I identified through BLAST, and included them in my alignment. I used the CLUSTAL W algorithm as implemented in MEGALIGN (DNA Star) to align my sequences with the published sequences. I further adjusted alignments by eye in MacClade version 4.06.

To check my assembled haplotype sequences for PCR artifacts such as chimeras, I scanned my alignments in both the Chimaera (Maynard Smith 1992, Posada and Crandall 2001, Posada 2002) and GENECONV (Padidam *et al.* 1999) applications in RDP version 2.0, recombination detection software (Martin *et al.* 2005). Potential PCR recombinant sequences were verified and removed from further analyses after visual inspection in MacClade version 4.06.

Bayesian analyses were conducted for each primer dataset using a GTR + G + I model of sequence evolution chosen by MrModelTest 2.2 in MrBayes (Ronquist and Huelsenbeck 2003). Four simultaneous MCMCMC chains were run for 10,000,000 generations sampling every 100 generations. Stationarity in likelihood scores was determined by plotting the -1nL against the generation. All trees below the observed stationarity level were discarded, resulting in a burnin of 75,000 generations. Estimation of best fit models for partial SSU rDNA gene sequences were performed using MrModelTest 2.2 (Nylander 2004).

Statistical Analyses

To compare between samples, estimated diversity was calculated using EstimateS version 8.0 (Colwell 2006), comparing number of clones sequenced to number of observed haplotypes based on my 99% assembling criterion. I also calculated the non-parametric richness estimator, Chao1, with EstimateS using 100 randomizations sampling

without replacement, along with Shannon's diversity index (H'). The rank abundance distribution of each dataset was used as the input to calculate each of these measures.

I performed Principal Coordinates Analysis (PCoA) and hierarchical clustering in UniFrac (Lupozone *et al.* 2006) with the Bayesian tree and a text file with sequence labels mapped to environmental samples as input. The UniFrac application measures the distance between two communities as that fraction of total the branch length in a tree that leads to descendants of members of either community, but not both (Lupozone *et al.* 2006). Using this approach, I can determine whether my communities sampled in different environments show relatedness to one another based on the environmental factors I measured. I applied weighted UniFrac with branch length normalization to my Bayesian tree. The distances were then plotted as points in a multidimensional space, one fewer than the number of samples. The principal coordinates describe how much of the variation each of the axes in this new space explains. These coordinates can then sometimes be correlated with environmental parameters of the samples.

I used the unweighted pair group method with arithmetic mean (UPGMA) hierarchical clustering algorithm, which clusters pairs of samples, and tested robustness of these clusters with jackknife analysis, a non-parametric estimator based on100 randomized sub-samples.

To determine the underlying distribution of diversity throughout my samples, I calculated curves for three models of community organization: log-series, log-normal and geometric series using the approach described by Magurran (2004). A log-series distribution represents a community that results from random assemblages of individuals taken from a larger community based on dispersal. I calculated the log-series distribution

using the equation $\alpha x^n/n$ where n is the abundance of haplotypes in the sample in this case, x is a fitted parameter, and α is Fisher's alpha, a diversity measure. I calculated the expected lognormal haplotype abundance distribution for each sample by calculating the mean and standard deviation of ln(abundance) and generating three expected abundance distributions for the species in the sample. I then calculated the mean abundance for each species, ranked from highest to lowest, and then calculated relative abundance.

A geometric series distribution is used to model communities organized based on species sorting by environment, assuming that the dominant species occupies some proportion of the resources of the environment k, the second most dominant then takes the same proportion k of the remainder, and this continues until every species is accommodated (Magurran 2004). I calculated the geometric series using the equation $n_i = NC_k k(1-k)^{i-1}$ where k = the proportion of niche space or resource that each species occupies, n_i = the total number of individuals in the ith species, N = the total number of individuals, and $C_k = [1-(1-k)^S]^{-1}$ and is a constant that ensures the $\sum n_i = N$.

For each distribution, I plotted expected along with the observed distributions of haplotypes. I generated these distributions using my defined haplotype as a sampling unit in place of the traditional unit of species. This unit is based on grouping together sequences of ≥99% similarity of my SSU rDNA marker. I also explored how the distribution models are affected by lowering the similarity cutoff to ≥98%, which reduces the haplotype diversity of the sample, to see if the observed patterns would change.

The observed distributions were compared to the model distributions using an Akaike goodness of fit test. The Akaike Information Criterion (AIC) was determined as the natural logarithm of the mean sum of squared deviations between observed and

predicted for all ranked haplotypes S plus an additional term to correct for the number of estimated parameters, k: (S+k)/(S-k-2). The lower the calculated AIC value, the better the fit of the data.

Results

Haplotypes Sequenced

In all, I sequenced 882 clones from the twelve samples and I obtained 67 ciliate haplotypes based on the 99% sequence similarity criterion (Table 3.2). Of the 67 haplotypes, 19 were identical to previously published environmental sequences, but only one matched the published sequence of a named morphospecies, *Strombidium biarmatum* (GB#AY541684) (Table 3.2). As in my previous work at sites along the Northwestern Atlantic (Doherty, *et al.* 2007), this haplotype was both abundant and ubiquitous throughout the samples, representing 480 of the 882 clones sequenced, and found at every station and depth. The proportion of common haplotypes (sampled $\geq$10 times) was 12 of the 67, or about 18%, while those sampled <10 times represented the remaining 82% of the sample. I found 26 singletons in the sample, representing 39% of the haplotypes.

I performed genealogical analyses of my SSU rDNA sequences along with sequences of published morphospecies from the Oligotrichia and Choreotrichia and uncultured environmental sequences from GenBank. The haplotypes I sequenced were fairly evenly distributed across the two subclasses with 27 haplotypes within the Choreotrichia and 40 haplotypes within the Oligotrichia (Table 3.2). However, the Oligotrichia sequenced were more numerically abundant in the samples than the

Choreotrichia (751 sequences versus 131), and were more likely to appear in more than one sample (30% singletons in Oligotrichia versus 52% singletons in Choreotrichia.

Estimates of diversity

To determine an estimate of the total diversity of ciliates in the samples, I calculated the nonparametric estimator Chao1, based on 100 randomizations, sampling without replacement in the program EstimateS version 8.0 (Colwell 2006) (Table 3.3). The estimator was used to assess total diversity of the sample by sub-sampling an input file of rank abundance data for each station and estimating diversity based on the abundance of haplotypes in the rare classes. The analysis predicts that I sampled most (>90%) of the diversity at only two stations (261 and 267), while at several other stations (260, 262, 263, 265) I captured less than 50% of the predicted overall diversity.

To compare my diversity with other ciliate studies, I calculated Shannon's diversity index (H') using the EstimateS program (Colwell 2006). My estimates of the Shannon index using molecular methods range from 0.9-2.29 (Table 3.3), which is higher than H' calculated for morphological samples of tintinnid ciliates in open ocean samples (~0.5-0.7) (Dolan *et al.* 2007), but similar to H' calculated for tintinnids sampled in coastal environments (~1.5-2.3) (Sitran *et al.* 2009).

Environmental Variation

The environmental parameters of salinity, temperature and oxygen concentration varied among the stations (Table 3.1). Salinity ranged from 21.2 to 29.6 over all samples, and during my sampling there was a visible front line caused by convergence of the fresher plume water with that of the open Sound. However, there was no apparent

relationship between any of these parameters and sample diversity as measured by Shannon's diversity index (H') (Table 3.2).

Although the environmental gradients were not correlated with any discernable changes in the ciliate assemblage, phytoplankton community composition and total abundance (as chlorophyll-a) did show some variations associated with the salinity gradient. For example, among the green algae prasinophytes were twice as abundant, as a fraction of total phytoplankton, outside the plume stations (Table 3.4); chlorophytes were more abundant within the plume, as were euglenophytes. Diatoms were the dominant taxon, comprising about 20% of the total phytoplankton chlorophyll, and their dominance was remarkably consistent across the whole sampling area.

As with salinity and temperature, the measures of phytoplankton community composition indicated by accessory pigments were not linked to ciliate diversity estimates. For instance, although stations 260, 262, and 263 had the highest levels of diversity as estimated by Chao1 (Table 3.2), the pigment concentrations varied widely from very high to very low at these three stations. The surface sample (0.25m) at station 260, closest to the river mouth, had the lowest pigment concentrations, about half of that recorded for the other stations, yet the ciliate diversity in this sample was not correspondingly low in accordance with the reduced concentration of prey (Table 3.2). Further, relative to the other samples, stations 258 and 259 contained elevated levels of phaeophytin, a product of the breakdown of chlorophyll that is typically associated with detritus in the water; this observation is consistent with the proximity of these two stations to shore, and the outflow from the Connecticut River.

Beta Diversity

I conducted analyses in Unifrac (Lozupone *et al.* 2006) to determine whether any significant clustering could be detected in the haplotype assemblages based on genealogical data. With the distances I obtained in my Bayesian genealogy, I conducted the analyses using my twelve samples as proxies for an environmental input. By treating each sample (depth and station) as an independent environment, I could determine patterning among the stations that my grosser measures could not detect. Principal coordinates analysis (PCoA) plots of the weighted UniFrac measurements show that the twelve environments fall into three fully supported groups (Fig. 3.2). Plotting the environments by the first two principal coordinates, which explain 42.75% and 39.83% of the variation respectively, reveals that samples 258 and 259 group together (cluster I), samples 264, 266, and 268 group together (cluster II), and the seven remaining samples (260, 261, 262, 263, 265, 267, and 269) group together (cluster III; Fig 3.2). UPGMA cluster analyses reveal full jackknife support for these three major groupings (Fig. 3.3).

Assessing diversity models

To compare distributions of haplotype diversity across the samples, I constructed rank abundance plots for each of the twelve samples, and pooled diversity data from each of the three clusters I detected in the Unifrac PCoA and UPGMA analyses. I implemented tests for goodness of fit for the three pooled clusters with each of three distributions (Table 3.5). None of the models fit my data particularly well as measured by the AIC scores. Cluster I, which grouped the two stations that were inside the plume front and nearest the shore, had the best fit to a log series distribution, while the other two clusters most closely fit a log normal distribution, albeit with weak scores. To determine the impact of *S. biarmatum,* the only taxon with high abundance at all sites, on the fit of

my data to modeled distributions, I evaluated the effect of its removal from the distribution. Without this species, lowest AIC values indicate that all three clusters most closely fit a log series distribution, though the fit for Cluster II (containing the lowest levels of diversity) is still poor (Table 3.5).

Discussion

As in my previous study (Doherty et al 2007), I found a single cosmopolitan and abundant species (*Strombidium biarmatum*) plus a large number of rare haplotypes. sOur survey in common with other published surveys, reveals a large amount of phylogenetic diversity in the Oligotrichia, where morphospecies sampling by molecular methods is sparse (Table 6; Doherty *et al.* 2007), indicating the importance of molecular tools in increasing my sampling capacity within these groups. In contrast, morphospecies with published sequences are more numerous in the Choreotrichia, probably due to their ease of identification using the extensive monographs of Kofoid and Campbell (1929, 1939) on the tintinnids, a group of choreotrichs with distinctive morphology in their outer shells or loricas (Dolan *et al.* 2007).

I observed no clear-cut relationship between the haplotype diversity within my samples and the environmental variables of depth, salinity, temperature, and phytoplankton pigments (Tables 3.1 and 3.3). While a subset of my samples has higher diversity as measured by Chao1 (Table 3.3), these higher diversity samples do not show any similarity across any of these variables (Tables 3.1 and 3.3). However, regression analysis of sample distance to shore versus sample diversity as measured by Shannon's diversity index (H'), does show a significant relationship ($p = 0.013$, $r^2 = 0.4762$). Distance from shore reflects increasing bottom depth, and increasing distance from the

mouth of the Connecticut River (Fig 3.1). With increasing distance from shore and the Connecticut River plume water, and with increasing water depth, diversity as measured by H' decreases.

Comparisons of clustering of samples based on phylogenetic diversity revealed three fully supported groups (Figs. 3.2 and 3.3). Similar to the results I obtained with diversity indices, these three clusters do not show any obvious relationships with the environmental parameters of temperature, salinity, and pigment concentrations. This may be due to the dynamic nature of the environment as the river plume front was moving away from shore and slightly westward while the tide was receding; hence, it was impossible to register environmental samples precisely with the DNA samples as the vessel drifted somewhat during sampling. In anticipation of this, I always sampled for DNA in relation to the visually observable front at the river plume edge, with samples taken either inside or outside the plume front. I am confident that the ciliate communities were collected across this gradient.

Based on the groupings revealed by the PCoA, I hypothesize that the main factors influencing composition of the ciliate assemblages in the samples were water depth, mixing, and proximity to the Connecticut River plume waters. Cluster I contains the two samples at the station closest to shore, with the shallowest ($\geq$5m) water depth, and greatest influence from the Connecticut River. The samples contain higher levels of phaeophytin, likely associated with detritus from the river plume, and also had the highest estimated contribution to the phytoplankton community from Euglenophytes, which are generally more abundant in freshwater. Additionally, this cluster differed substantially from the other two clusters in the distribution of Oligotrichia and

Choreotrichia haplotypes (Table 3.6). Choreotrichia and Oligotrichia are represented in fairly equivalent proportions in cluster I (40:60 Choreotrich to Oligotrich), in contrast to clusters II and III, where Oligotrichia dominate (Table 3.6).

Cluster II is distinctive in that it represents three surface samples ($\simeq$0.5m) over deeper water (10-40m) and at a greater distance from shore than the Cluster I samples. At this bathymetric depth, the water column experiences less mixing than in the near-shore environment of Cluster I, and may thus contain a ciliate assemblage characteristic of the open coast. Cluster II is predominantly characterized by ciliates in the Oligotrichia, representing all but 4 of 176 sequences (Table 3.6), and, in general, some of the lowest levels of diversity estimated by Chao1 (Table 3.3).

Cluster III, containing the largest portion of the samples, represents all of the non-surface water samples, and most of the samples taken in shallower areas along the plume front. I predict that Cluster III may represent a well-mixed environment, where similar communities exist in both the shallow and deep waters, and the coastal signal is weakest. Cluster III contains a greater proportion of ciliates in the Oligotrichia than the Choreotrichia (69% versus 31%), but less extremely so as compared to Cluster II (Table 3.6). This cluster also contains a greater range of estimated diversity, from the least diverse sample (267), to the samples with the highest estimated diversity (262, 263, 265, and 260) (Table 3.3).

Considering that environmental differences potentially structure ciliate assemblages in this coastal setting, I also wanted to test the hypothesis that the assembly of ciliate communities is consistent with expectations under neutrality (Hubbell 2001). In a survey of tintinnid communities (Choreotrichia) in the South Pacific Ocean, Dolan *et*

al. (2007) determined that the assemblages at all but one of the 21 stations they sampled fit predictions of neutral community theory in that distributions best fit a log series distribution. Neutral theory predicts that communities are assembled by stochastic factors such as immigration, death and birth, rather than being structured by environmental conditions (Hubbell 2001, Dolan *et al.* 2007). In contrast, an investigation on tintinnid communities in the coastal Mediterranean Sea revealed distribution patterns with greater similarity to a modeled log normal distribution, indicating a stronger impact of the environment in structuring ciliate assemblages in this setting (Sitran *et al.* 2009).

I assessed how haplotype data obtained using molecular approaches compare to these findings. Because my Unifrac analysis revealed well-supported groupings, I constructed my haplotype distributions based on these three clusters (I-III; Fig. 3.2). When I included all data, the fits to models were poor as measured by AIC in clusters II and III, with these two clusters matching a log normal distribution, while I get a better fit in cluster I to a log series distribution (Table 3.5). I then removed the ubiquitous and abundant *Strombidium biarmatum*. *S. biarmatum* is a very small species that may have a unique biology compared to the remaining haplotypes, and I wanted to determine its impact on the distributions. Without *S. biarmatum*, all clusters fit the log series distribution best, though the fit for cluster II remains poor (Table 3.5). This suggests the possibility that the signal from *S. biarmatum* is strong enough to mask the signal from the remainder of the community, which may be assembled in a manner consistent with neutrality.

These analyses, and others (Pedros-Alio 2006, Sogin *et al.* 2006), suggest that the molecular pattern of microbial taxa can be overwhelmed by a small number of

numerically abundant taxa that dominate samples -- in my case *Strombidium biarmatum.* Increased molecular sampling effort in many microbial communities shows increasing diversity of the rare community members (Pedros-Alio 2006, Sogin *et al.* 2006, Massana and Pedros-Alio 2008). In the case of ciliates, molecular surveys from Long Island Sound waters using my primers captured 27 ciliate haplotypes from a total of 229 clones in 2005 (Doherty *et al.* 2007), and this study, with more intensive sampling (882 clones), found an additional 62 new haplotypes not captured in the previous survey, with only 5 haplotypes of the original 27 recurring in the present study.

Finally, the unit of diversity I are employing in my molecular samples to calculate my distribution models, is not clearly linked to a taxonomic identity. Because so few ciliate morphospecies have been characterized molecularly, I am required to use percent sequence similarity as a proxy for taxonomic difference. I defined a haplotype as ciliate sequences with ≥99% similarity in my dataset. How this unit of diversity compares with the unit morphospecies such as described in Dolan *et al.* (2007) has yet to be revealed. I did however investigate the impact of reducing the stringency of this cutoff to ≥98% and I predicted that reducing the stringency would not influence the calculated distributions since the effect of lowering the cutoff only reduces the number of rare haplotypes in the sample. I tested this effect on the sample where *Strombidium biarmatum* had been removed as well. In summary, the effect of changing the haplotype definition to ≥98%, with or without *Strombidium biarmatum,* produced a very small effect in Clusters I-III, in every case showing only marginal difference between modeled log normal and log series distributions (data not shown).

Data from other protist molecular surveys in ocean environments show a distinct structuring of haplotype diversity that is associated with environmental differences on a larger scale than I were examining (Countway *et al.* 2007, Not *et al.* 2007). This has also been found in prokaryotic communities (Fuhrman *et al.* 2006) in marine planktonic systems, as well as fresh water stream ecosystems (Crump *et al.* 2007). I detected differences in community composition, diversity, and haplotype distributions between my near shore samples (Cluster I) and my deeper water samples (Clusters II and III), and between surface samples (Cluster II) and deeper/mixed samples (Cluster III). I also observed significantly greater diversity (Shannon's) associated with proximity to the river plume (Fig. 3.1). Based on these results, it is clear that coastal ciliate assemblages differ markedly over very short (km) spatial scales, but this signal is difficult to correlate with the measured environmental parameters I investigated.

Table 3.1 Collection data for samples collected near CT River Plume on Long Island Sound.

Sample	Depth(m)	Lat (° N)	Long (°W)	Temp(°C)	Salinity	Oxygen
258	1	41 15.49	072 19.68	13.53	28.83	8.46
259	2	41 15.48	072 19.66	13.54	29.63	8.48
260	0.25	41 15.47	072 19.71	14.5	26.43	7.67
261	4	41 15.46	072 19.82	13	27.97	7.4
262	0.5	41 15.27	072 20.23	14.65	25.42	7.32
263	4	41 15.13	072 19.40	12.89	27.95	7.45
264	0.5	41 14.85	072 19.60	13.75	27.12	7.68
265	4	41 14.85	072 19.46	12.78	28.17	7.28
266	0.5	41 14.77	072 49.58	16.08	21.18	7.79
267	4	41 14.76	072 19.46	13.28	27.82	7.04
268	0.5	41 14.25	072 19.48	15.76	26.9	n/a
269	4	41 14.06	072 19.45	14	27.56	n/a

Table 3.2 Haplotype diversity of the twelve samples from Long Island Sound..

	Sample and unifrac cluster												
	I	I	III	III	III	III	II	III	II	III	II	III	
Sequence Name / GenBank #	**258**	**259**	**260**	**261**	**262**	**263**	**264**	**265**	**266**	**267**	**268**	**269**	Clade
Strombidium biarmatum AY541684	12	17	60	50	38	56	39	40	25	57	30	56	O
EF553452	7	3	6				21	6	18	5	15	1	O
EF553411	6	2		6	11		1	9					O
258_02	10	13											O
EF553405			1				1	4	3	3			O
EF553454				2	2	1		1	2			3	O
EF553421				2		4		4					O
EF553415			1		3		1		3			1	O
EF553413		1								2		5	O
hbp84			1		1	4							O
260_07			1		2	2		1					O
260_08		1	1							1		1	O
262_11		2			1								O
258_04	2								1				O
263_01						3							O
262_02					1	1							O
hbp86		1						1					O
258_07	1				1								O
260_01			1		1								O
260_09			1					1					O
hbp95	14	20	1										C
EF553456	8	2	4	2	4		4	3	4				C
EF553396	3		3	3		1		2					C
258_05	3			2		1		1		3	1		C
260_05			5		5	1							C
260_06			1									5	C
262_14		1		4	1								C
EF553416					1		1		2		1		C
258_06	3							1					C
EF553401			1	2									C
262_12					1					2			C
262_13					1	1						1	C
hbp88				2	1								C
262_01					3								C
259_01		2											C
259_02		1				1							C
266_01		1							1				C
Unique haplotypes	4	3	2	3	8	4		4	1		1		C/O

Table 3.3 Measured and estimated diversity of Long Island Sound samples by station number.

Sample	Haplotypes	Chao1 Estimate (95% Confidence)		% Div sampled	H'
258	15	17.25	(2.2, 20.75)	87.0	2.4
259	17	27.13	(-7.87, 35.15)	62.7	2.1
260	17	83	(43.89, 130.99)	20.5	1.46
261	13	13.75	(0.99, 8.16)	94.5	1.49
262	26	74.17	(36.43, 142.77)	35.1	2.29
263	16	76.5	(38.91, 226.73)	20.9	1.34
264	7	13	(4.68, 29.93)	53.8	1.1
265	18	43	(19.93, 98.28)	41.9	1.94
266	10	12.25	(1.98, 16.79)	81.6	1.64
267	7	7.25	(0.41, 6.05)	96.6	0.9
268	5	8	(2.25, 19.26)	62.5	0.9
269	8	14	(5.02, 34.4)	57.1	0.94

Table 3.4 Proportion of the total phytoplankton community (as chlorophyll a) contributed by various groups of algae for each sample (258-269).

Sample		Diatom	Dinoflagellate	Cyanobacteria	Prasinophyte	Chlorophyte	Cryptophyte	Prymnesiophyte	Raphidophyte	Chrysophyte	Euglenophyte
258	**I**	0.16	0.02	0.05	0.00	0.05	0.34	0.16	0.12	0.00	0.10
259	**I**	0.20	0.03	0.05	0.04	0.04	0.35	0.11	0.10	0.00	0.08
262	**I**	0.09	0.01	0.14	0.00	0.15	0.16	0.25	0.08	0.05	0.06
263	**I**	0.25	0.03	0.05	0.15	0.00	0.40	0.03	0.04	0.01	0.04
266	**I**	0.17	0.03	0.08	0.10	0.06	0.28	0.05	0.10	0.07	0.06
267	**I**	0.23	0.02	0.05	0.22	0.00	0.39	0.01	0.02	0.05	0.01
	mean	0.18	0.02	0.07	0.08	0.05	0.32	0.10	0.08	0.03	0.06
	s.e.	0.02	0.00	0.02	0.04	0.02	0.04	0.04	0.02	0.01	0.01
260	**O**	0.20	0.03	0.05	0.21	0.00	0.35	0.01	0.06	0.08	0.02
261	**O**	0.25	0.03	0.05	0.21	0.00	0.38	0.02	0.03	0.04	0.00
264	**O**	0.19	0.03	0.05	0.22	0.00	0.36	0.01	0.05	0.07	0.01
265	**O**	0.23	0.02	0.06	0.21	0.00	0.41	0.01	0.02	0.03	0.01
268	**O**	0.11	0.03	0.12	0.25	0.00	0.26	0.04	0.07	0.11	0.01
269	**O**	0.18	0.04	0.08	0.24	0.00	0.31	0.03	0.06	0.06	0.01
	mean	0.19	0.03	0.07	0.22	0.00	0.34	0.02	0.05	0.06	0.01
	s.e.	0.02	0.00	0.01	0.01	0.00	0.02	0.00	0.01	0.01	0.00

Table 3.5 Results of analysis of haplotype abundance distributions.

Haplotype Assembly	Cluster	Log Series	Geometric	Log Normal
99% with *S. biarmatum*				
	I	**-0.226**	1.807	0.202
	II	1.747	4.631	**1.544**
	III	0.186	7.759	**0.167**
99% without *S. biarmatum*				
	I	**-0.860**	3.907	0.306
	II	**0.832**	1.059	1.270
	III	**-1.606**	2.887	1.121

Table 3.6 Distribution of ciliates in the subclasses Choreotrichia and Oligiotrichia by Unifrac cluster.

	I	II	III
Sequences			
Choreotrich	59	4	68
Oligotrich	87	172	492
Total	146	176	560
% Choreotrich	40	2	12
% Oligotrich	60	98	88
Haplotypes			
Choreotrich	13	4	32
Oligotrich	19	18	72
Total	32	22	104
% Choreotrich	41	18	31
% Oligotrich	59	82	69

.

Figure 3.1 Chart of study area near the mouth of the Connecticut River. Locations of the surface stations are indicated, along with approximate positions of the river plume front (dashed lines), as indicated by the visually-observable surface slick, as well as the backscatter signal and currents from an acoustic doppler current profiler (ADCP). Current and backscatter measurements were made along the two tracks indicated by red and green lines. The front moved slightly westward and southward with the receding tide during the sampling.

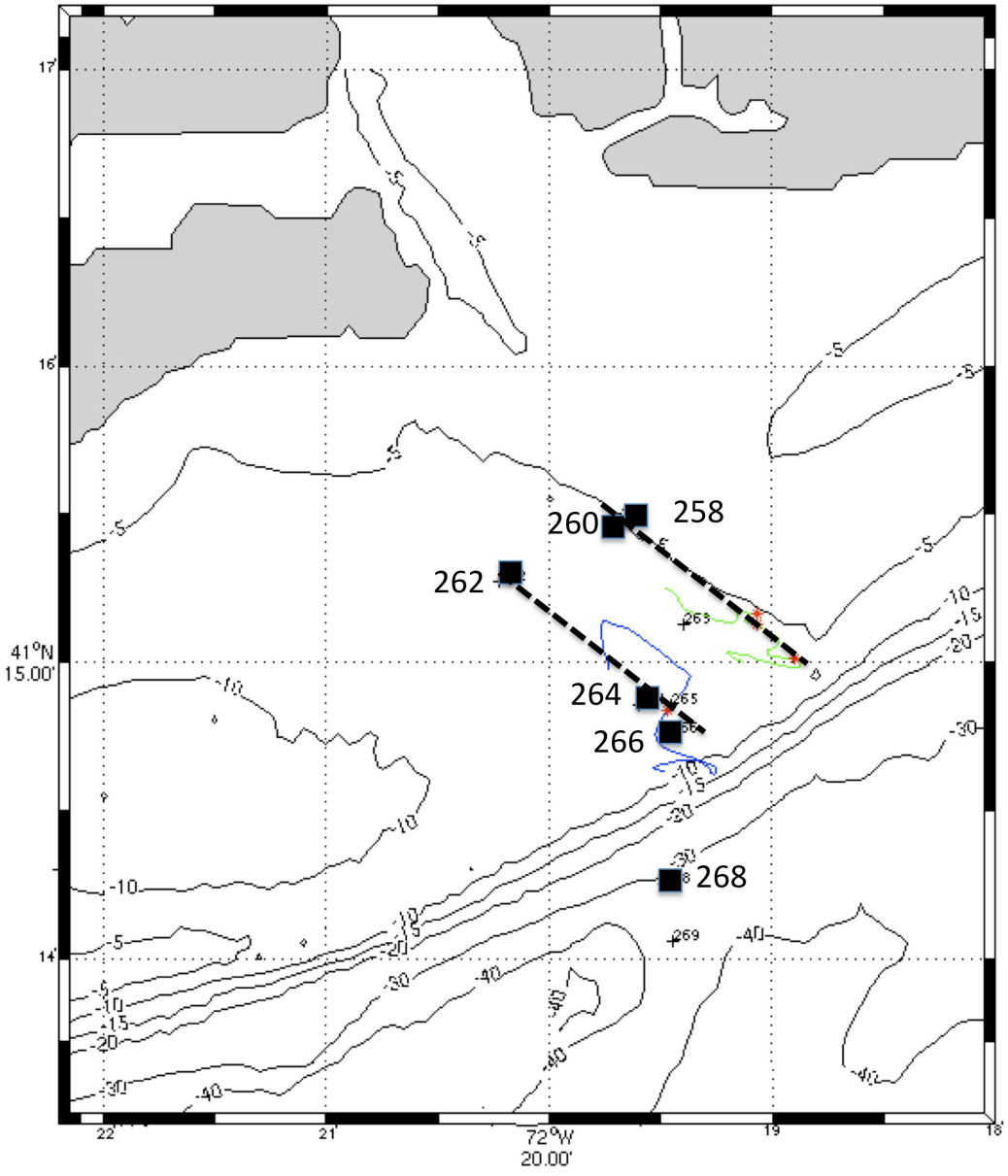


Figure 3.2 Results for Principle Coordinates Analysis (PCoA) based on branch lengths in the Bayesian tree and environmental data for the twelve stations sampled from Long Island Sound.

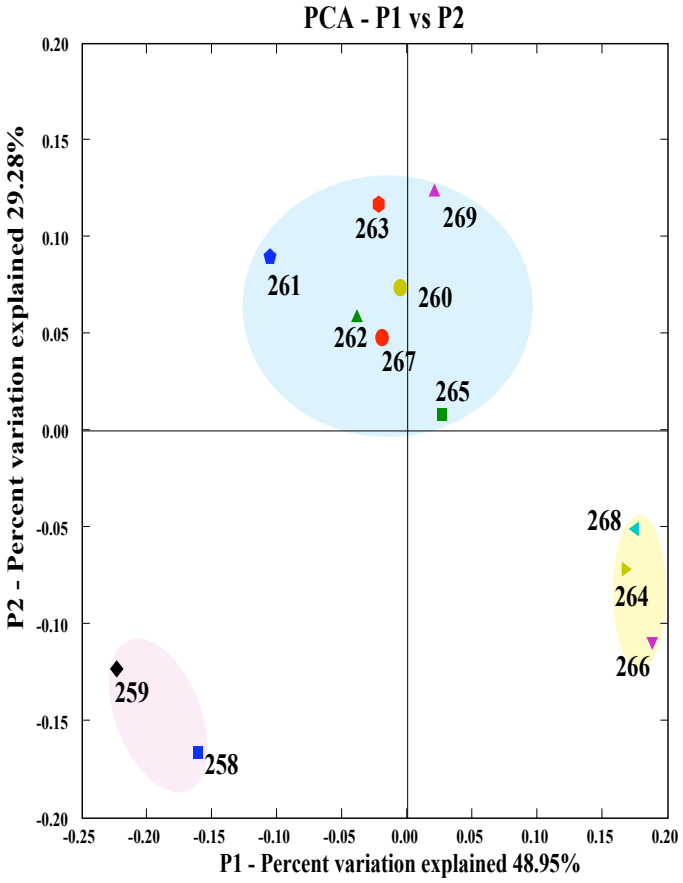
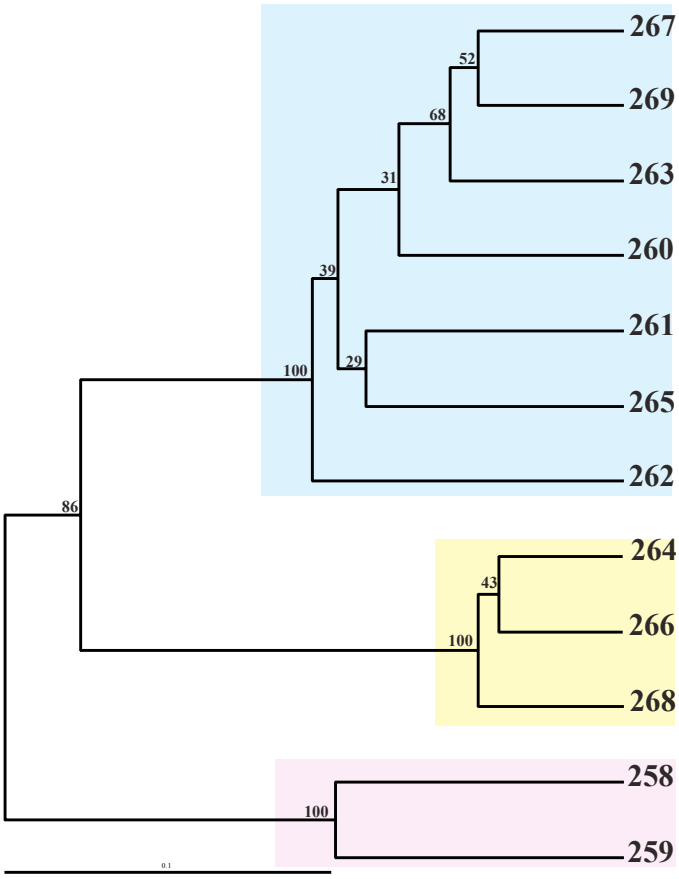


Figure 3.3 Results from hierarchical clustering using UPGMA for the twelve samples with jackknife support from 100 replicates.



BIBLIOGRAPHY

Agatha, S. (2004) A cladistic approach for the classification of Oligotrichid ciliates (Ciliophora: Spirotrichia). *Acta Protozoologica* 43: 201-217.

Agatha, S., Struder-Kypke, M.C., Beran, A., Lynn, D.H. (2005) *Pelagostrobilidium neptuni* (Montagnes and Taylor, 1994) and *Strombidium biarmatum* nov. spec. (Ciliophora, Oligotrichea): phylogenetic position inferred from morphology, ontogenesis, and gene sequence data. *European Journal of Protistology*. 41: 65-83.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25: 3389–3402.

Ausubel F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (1992) *Short Protocols in Molecular biology: a Compendium of methods from Current Protocols in Molecular Biology*. Wiley-Liss, New York.

Behnke, A., Bunge, J., Barger, K., Breiner, H.W., Alla,V., Stoeck, T. (2006) Microeukaryote community patterns along an O-2/H2S gradient in a supersulfidic anoxic Fjord (Framvaren, Norway). *Applied And Environmental Microbiology* 72: 3626-3636.

Bell, G. (2001) Neutral Macroecology. *Science* 293: 2413-2418.

Belmonte, G., Miglietta, A., Rubino, F., Boero F. (1997) Morphological convergence of resting stages of planktonic organisms: a review. *Hydrobiologia* 355: 159-165.

Calbet, A. and Saiz, E. (2005) The ciliate-copepod link in marine ecosystems. *Aquatic Microbial Ecology* 38: 157-167.

Capriulo, G.M. and Carpenter, E.J. (1983) Abundance, species composition and feeding impact of tintinnid micro-zooplankton in central Long Island Sound. *Marine Ecolology Progress Series* 10: 277-288.

Capriulo, G.M., Smith, G., Troy, R., Wikfors, G.H., Pellet, J., Yarish, C. (2002) The planktonic food web structure of a temperate zone estuary and its alteration due to eutrophication. *Hydrobiologia* 475: 263-333.

Caron, D.A., Countway, P.D., Brown, M.V. (2004) The growing contributions of molecular biology and immunology to protistan ecology: molecular signatures as ecological tools. *Journal Of Eukaryotic Microbiology* 51: 38-48.

Caudill, C. C. and Bucklin, A. (2004) Molecular phylogeography and evolutionary history of the estuarine copepod, *Acartia tonsa*, on the Northwest Atlantic coast. *Hydrobiologia* 511: 01-102

Colwell, R.K. (2005) EstimateS: Statistical estimation of species richness and shared species from samples.

Colwell, R. K. (2006) *EstimateS*: Statistical estimation of species richness and shared species from samples. Version 8. Persistent URL <purl.oclc.org/estimates>

Cordova-Kreylos, A.L., Cao, Y.P., Green, P.G., Hwang, H.M., Kuivila, K.M., LaMontagne, M.G., Van De Werfhorst, L.C., Holden, P.A., Scow, K.M. (2006) Diversity, composition, and geographical distribution of microbial communities in California salt marsh sediments. *Applied And Environmental Microbiology* 72: 3357-3366.

Corliss, J.O. and Esser, S. C. (1974) Comments on the role of the cyst in the life cycle and survival of free-living protozoa. *Transactions of the American Microscopical Society.* 93: 578-593.

Costas, B., McManus, G., Doherty, M., Katz, L. (2007) Use of species-specific primers and PCR to measure the distributions of planktonic ciliates in coastal waters. *Limnology and Oceanography: Methods* 5: 163-173.

Countway, P. D., Gast, R., Dennett, M. R., Savai, P., Rose, J. M., Caron, D. A. (2007) Distinct protistan assemblages characterize the euphotic zone and deep sea (2500m) of the western North Atlantic (Sargasso Sea and Gulf Stream). *Environmental Microbiololgy* 9: 1219-1232.

Countway, P.D., Gast, R.J., Savai, P., Caron, D.A. (2005) Protistan diversity estimates based on 18S rDNA from seawater incubations in the western North Atlantic. *Journal Of Eukaryotic Microbiology* 52: 95-106.

Crump, B.C., Adams, H.E., Hobbie, J. E., and Kling, G.W. (2007) Biogeography of bacterioplankton in lakes and streams of an arctic tundra catchment *Ecology* 88: 1365-1378.

Cuvelier, M. L., Ortiz, A., Kim, E., Moehlig, H., Richardson, D. E., Heidelberg, J.F., Archibald, J. M., Worden, A.Z. (2008) Widespread distribution of a unique marine protistan lineage. *Environmental Microbiology* 10: 1621-1634.

Dawson, S.C. and Pace, N.R. (2002) Novel kingdom-level eukaryotic diversity in anoxic environments. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 99: 8324-8329.

DeLong, E.E. (2005) Microbial community genomics in the ocean. *Nature Reviews Microbiology* 3: 459-469.

De Stasio, B. T. (1989) The seed bank of a freshwater crustacean: copepodology for the plant ecologist. *Ecology* 70: 1377-1389.

Diez, B., Pedros-Alio, C., Massana, R. (2001) Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. *Applied And Environmental Microbiology* 67: 2932-2941.

Doherty, M., Costas, B.A., McManus, G.B., Katz, L.A. (2007) Culture-independent assessment of planktonic ciliate diversity in coastal northwest Atlantic waters. *Aquatic Microbial Ecology* 48: 141-154.

Dolan, J.R., Lemee, R., Gasparini, S., Mousseau, L., Heyndrickx, C. (2006) Probing diversity in the plankton: using patterns in Tintinnids (planktonic marine ciliates) to identify mechanisms. *Hydrobiologia* 555: 143-157.

Dolan, J.R., Ritchie, M.E., Ras, J. (2007) The "neutral" community structure of planktonic herbivores, tintinnid ciliates of the microzooplankton, across the SE Tropical Pacific Ocean. *Biogeoscience* 4: 297-310.

Edgcomb, V.P., Kysela, D.T., Teske, A., Gomez, A.D., Sogin, M.L. (2002) Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 99: 7658-7662.

Esteban, G. F. and Finlay B. J. (2003) Cryptic freshwater ciliates in a hypersaline lagoon. *Protist* 154: 411-418.

Esteban, G. F. and Finlay, B. J. (2007) Exceptional species richness of ciliated protozoa in pristine intertidal rock pools. *Marine Ecology Progress Series* 335: 133-141.

Fenchel, T. (1969) The ecology of marine microbenthos IV. Structure and function of the benthic ecosystem, its chemical and physical factors and the microfauna communities with special reference to the ciliated protozoa. *Ophelia* 6: 1-182.

Fenchel,T. and Jonsson, P.R. (1988) The functional biology of *Strombidium sulcatum*, a marine oligotrich ciliate (Ciliophora, Oligotrichina). *Marine Ecology Progress Series* 48: 1-15.

Fileman, E.S. and Leakey, R.J.G. (2005) Microzooplankton dynamics during the development of the spring bloom in the north-east Atlantic. *Journal Of The Marine Biological Association Of The United Kingdom* 85: 741-753.

Finlay, B.J. (2002) Global dispersal of free-living microbial eukaryote species. *Science* 296: 1061-1063.

Foissner, W., Muller, H., Agatha S. (2007) A comparative fine structural and phylogenetic analysis of resting cysts in oligotrich and hypotrich Spirotrichea (Ciliophora). *European Journal of Protistology*. 43: 295-314.

Fuhrman, J.A., Hewson, I., Schwalbach, M.S., Steele, J.A., Brown, M.V., Naeem, S. (2006) Annually reoccurring bacterial communities are predictable from ocean conditions. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 103: 13104-13109.

Fuhrman, J.A., Steele, J.A., Hewson, I., Schwalbach, M.S., Brown, M.V., Green, J.L., and Brown, J.H. (2008) A latitudinal diversity gradient in planktonic marine bacteria. *Proceedings of the National Academy of Sciences U S A* 105: 7774-7778.

Giovannoni, S.J. and Stingl, U. (2005) Molecular diversity and ecology of microbial plankton. *Nature* 437: 343-348.

Hairston N. G., Kearns, C. M., and Ellner, S. P. (1996). Phenotypic variation in a zooplankton egg bank. *Ecology* 77: 2382-2392.

Hamels, I., Muylaert, K., Sabbe, K., Vyverman, W. (2005) contrasting dynamics of ciliate communities in sandy and silty sediments of an estuarine intertidal flat. *European Journal of Protistology* 41: 241-450.

Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography* Princeton University Press, Princeton, New Jersey, USA.

Hubbell, S.P. (2005) Neutral theory in community ecology and the hypothesis of functional equivalence. *Functional Ecology* 19: 166-172.

Hubbell, S.P. (2006) Neutral theory and the evolution of ecological equivalence. *Ecology* 87: 1387-1398.

Hughes, J. B. and Hellman, J. J. (2005) The application of rarefaction techniques to molecular inventories of microbial diversity. *Methods in Enzymology* 397: 292-308.

Jonsson, P.R., Johnsson, M., Pierce, R.W. (2004) Attachment to suspended particles may improve foraging and reduce predation risk for tintinnid ciliates. *Limnology and Oceanography* 49: 1907-1914.

Kamiyama, T. (1994) Effects of extracellular products from phytoplankton on the excystment of tintinnids from marine sediments. *Marine Ecology Progress Series* 105: 199-201.

Kamiyama, T. and Aizawa, Y. (1992) Effects of termperature and light on tintinnid excystment from marine sediments. *Nippon Suisan Gakkaishi* 58: 877-884.

Kamiyama, T. and Matsuyama, Y. (2005) Temporal changes in the ciliate assemblage and consecutive estimates of their grazing effect during the course of a *Heterocapsa circularisquama* bloom. *Journal Of Plankton Research* 27: 503-503.

Katz, L.A., McManus, G.B., Snoeyenbos-West, O.L.O., Griffin, A., Pirog, K., Costas, B., Foissner, W. (2005) Reframing the 'everything is everywhere' debate: evidence for high gene flow and diversity in ciliate morphospecies. *Aquatic Microbial Ecology* 41: 55-65.

Kim, Y.O. and Taniguchi, A. (1995) Excystment of the oligotrich ciliate *Strombidium conicum*. *Aquatic Microbial Ecology*. 9: 149-156.

Kim, Y. O. and Taniguchi, A. (1997) Seasonal variation of excystment pattern of the planktonic oligotrich ciliate *Strombidium conicum*. *Marine Biology* 128: 207-212.

Ley, R.E., Harris, J.K., Wilcox, J., Spear, J.R., Miller, S.R., Bebout, B.M., Maresca, J.A., Bryant, D.A., Sogin, M.L., Pace, N.R. (2006) Unexpected diversity and complexity of the Guerrero Negro hypersaline microbial mat. *Applied And Environmental Microbiology* 72: 3685-3695.

Lopez-Garcia, P., Philippe, H., Gail, F., Moreira, D. (2003) Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the Mid-Atlantic Ridge. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 100: 697-702.

Lopez-Garcia, P., Rodriguez-Valera, F., Pedros-Alio, C., Moreira, D. (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* 409: 603-607.

Lozupone, C.A., Hamady, M., Kelley, S. T., Knight, R. (2007) Quantitative and qualitative ß diversity measures lead to different insights into factors that structure microbial communities. *Applied and Environmental Microbiology* 73: 1576-1585.

Lozupone, C., Hamady, M., Knight, R. (2006) UniFrac – an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* 7: 371.

Lozupone, C. A. and Knight, R. (2007) Global patterns in bacterial diversity. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 104: 11436-11440.

Mackey, M. D., Mackey, D. J., Higgins, H. W., Wright, S. W. (1996) CHEMTAX – A program for estimating class abundances from chemical markers: Application to HPLC measurements of phytoplankton. *Marine Ecology Progress Series* 144: 265-283.

Maddison, W. P. and D. R. Maddison. 1992. MacClade: analysis of phylogeny and character evolution. Sinauer, Sunderland, Mass.

Madoni, P. (2006) Benthic ciliates in Adriatic Sea lagoons. *European Journal of Protistology* 42: 165-173.

Maeda, M. (1986) An illustrated guide to the species of the families Halteriidae and Strobilidiidae (Oligotrichida, Ciliophora) free swimming protozoa common in the aquatic environment. *Bulletin of the Ocean Research Institute* 21: 1-67.

Maeda, M. and Carey, P.G. (1985) An illustrated guide to the species of the family Strombidiidae (Oligotrichida, Ciliophora), free swimming protozoa common in the aquatic environment. *Bulletin of the Ocean Research Institute, University of Tokyo* 19: 1-68.

Magurran, A. E. *Measuring Biological Diversity* London, U.K., Blackwell Publishing, 2004.

Marcus, N. H., Lutz, R., Burnett, W., Cable, P. (1994). Age, viability, and vertical distribution of zooplankton resting eggs from an anoxic basin: evidence of an egg bank. *Limnology and Oceanography* 39: 154-158.

Marshall, S.M. (1969) Protozoa. Order: Tintinnida. Conseil *International pour l'Exploration de la Mer*, Zooplankton Sheets 117-127: 1-81.

Martin, A.P. (2002) Phylogenetic approaches for describing and comparing the diversity of microbial communities. *Applied and Environmental Microbiology* 68: 3673-3682.

Martin, D.P., Williamson, C., Posada, D. (2005) RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics* 21: 260-262.

Massana, R. and Pedros-Alio, C. (2008) Unveiling new microbes in the surface ocean. *Current Opinion in Microbiology* 11: 213-218.

Maynard Smith, J. (1992) Analysing the mosaic structure of genes. *Journal of Molecular Evolution* 34: 126-129.

McGill, B.J., Etienne, R.S., Gray, J.S., Alonso, D., Anderson, M.J., Benecha, H.K., Dornelas, M., Enquist, B.J., Green, J.L., Fangliang, H., Hurlbert, A.H., Magurran, A.E., Marquet, P.A., Maurer, B.A., Ostling, A., Soykan, C.U., Ugland, K.I., White, E.P. (2007) Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters* 10: 995-1015.

McManus, G.B. and Fuhrman, J.A. (1986) Photosynthetic pigments in the ciliate *Laboea strobila* from Long Island Sound, USA. *Journal of Plankton Research* 8: 317-327.

Modigh, M. (2001) Seasonal variations of photosynthetic ciliates at a Mediterranean coastal site. *Aquatic Microbial Ecology* 23: 163-175.

Modigh, M., Castaldo, S., Saggiomo, M., Santarpia, I. (2003) Distribution of tintinnid species from 42 degrees N to 43 degrees S through the Indian Ocean. *Hydrobiologia* 503: 251-262.

Montagnes, D.J.S., Lowe, C.D., Poulton, A., Jonsson, P.R. (2002a) Redescription of of *Strombidium oculatum* Gruber 1884 (Ciliophora, Oligotrichia). *Journal of Eukaryotic Microbiology* 49: 329-337.

Montagnes, D.J.S. and Lynn, D.H. (1991) Taxonomy of choreotrichs, the major marine ciliates, with emphasis on the aloricate forms. *Marine Microbial Food Webs* 5: 59-74.

Montagnes, D.J.S., Wilson, D., Brooks, S.J., Lowe, C., Campey, M. (2002b) Cyclical behavior of the tide-pool ciliate *Strombidium oculatum*. *Aquatic Microbial Ecology* 28: 55-68.

Moon-van der Staay, S.Y., De Wachter, R., Vaulot, D. (2001) Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* 409: 607-610.

Muller, H. (2000) Evidence of dormancy in planktonic oligotrich ciliates. *Verhandlungen des Internationalen Verein Limnologie* 27: 3206-3209.

Muller, H. (2002) Laboratory study of the life cycle of a freshwater strombidiid ciliate. *Aquatic Microbial Ecology* 29: 189-197.

Muller, H. (2007) Live observation of excystment in the spirotrich ciliate *Meseres corlissi*. *European Journal of Protistology* 43: 95-100.

Not, F., Gausling, R., Azam, F., Heidelberg, J.F., Worden, A.Z. (2007) Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environmental Microbiology* 9: 1233-1252.

Nylander, J. (2004) MrModeltest v2. Evolutionary Biology Centre, Uppsala University; program distributed by the author.

Padidam, M., Sawyer S., Fauquet, C.M. (1999) Possible emergence of new geminiviruses by frequent recombination. *Virology* 265: 218-225.

Paranjape, M.A. (1980) Occurrence and significance of resting cysts in a hyaline tintinnid, *Helicostmella subulata* (Ehre.) Jorgensesn. *Journal of Experimental Marine Biology and Ecology*. 48: 23-33.

Patterson, D. J., Larsen, J., Corliss, J. O. (1989) The ecology of heterotrophic flagellates and ciliates living in marine sediments. *Progress in Protistology* 3: 185-277.

Pedros-Alio, C. (2006) Marine microbial diversity: can it be determined? *Trends in Microbiology* 14: 257-263.

Pierce, R.W., Turner, J.T. (1992) Ecology of planktonic ciliates in marine food webs. *Reviews In Aquatic Sciences* 6: 139-181.

Posada, D. (2002) Evaluation of methods for detecting recombination from DNA sequences: Empirical data. *Molecular Biology And Evolution* 19: 708-717.

Posada, D. and Crandall, K.A. (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology and Evolution* 16: 37-45.

Reid, P.C. (1987) Mass encystment of a planktonic oligotrich ciliate. *Marine Biology* 95: 221-230.

Reid, P. C. and John, A.W. G. (1978) Tintinnid cysts. *Journal of the Marine Biological Association of the United Kingdom* 58: 551-557.

Richards, T.A. and Bass, D. (2005) Molecular screening of free-living microbial eukaryotes: diversity and distribution using a meta-analysis. *Current Opinion In Microbiology* 8: 240-252.

Richards, T.A., Vepritskiy, A.A., Gouliamova, D.E., Nierzwicki-Bauer, S.A. (2005) The molecular diversity of freshwater picoeukaryotes from an oligotrophic lake reveals diverse, distinctive and globally dispersed lineages. *Environmental Microbiology* 7: 1413-1425.

Romari, K. and Vaulot, D. (2004) Composition and temporal variability of picoeukaryote communities at a coastal site of the English Channel from 18S rDNA sequences. *Limnology And Oceanography* 49: 784-798.

Ronquist, F. and Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572-1574.

Rosetta, C.H. and McManus, G.B. (2003) Feeding by ciliates on two harmful algal bloom species, *Prymnesium parvum* and *Prorocentrum minimum*. *Harmful Algae* 2: 109-126.

Sherr, E.B., Sherr, B.F., Fallon, R.D., Newell, S.Y. (1986) Small, aloricate ciliates as a major component of the marine heterotrophic nanoplankton. *Limnology And Oceanography* 31: 177-183.

Shimeta, J., Amos, C. L., Beaulieu, S. E., Ashiru, O.M. (2002) Sequential resuspension of protists by acceleration tidal flow: implication for community structure in the benthic boundary layer. *Limnological Oceanography* 47: 1152-1164.

Shimeta, J., Gast, R. J., Rose, J.M. (2007) Community structure of marine sedimentary protists in relation to flow and grain size. *Aquatic Microbial Ecology* 48: 91-104.

Shimeta, J. and Sisson, J.D. (1999) Taxon-specific tidal resuspension of protists into the subtidal benthic boundary layer of a coastal embayment. *Marine Ecology Progress Series* 177: 51-62.

Sitran, R., Bergamasco, A., Decembrini, F., Guglielmo, L. (2009) Microzooplankton (tintinnid ciliates) diversity: coastal community structure and driving mechanisms in the southern Tyrrhenian Sea (Western Mediterranean). *Journal of Plankton Research* 31: 153-170.

Sloan, W. T., Woodcock, S., Lunn, M., Head, I. M., Curtis, T. P. (2007) Modeling taxa-abundance distributions in microbial communities using environmental sequence data. *Microbial Ecology* 53: 443-455.

Snoeyenbos-West, O.L.O., Salcedo, T., McManus, G.B., Katz, L.A. (2002) Insights into the diversity of choreotrich and oligotrich ciliates (Class: Spirotrichea) based on genealogical analyses of multiple loci. *International Journal Of Systematic And Evolutionary Microbiology* 52: 1901-1913.

Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R., Arrieta, J. M., Herndl, G. J. (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings Of The National Academy Of Sciences Of The United States Of America* 103: 12115-12120.

Stoeck, T., Hayward, B., Taylor, G.T., Varela, R., Epstein, S.S. (2006) A multiple PCR-primer approach to access the microeukaryotic diversity in environmental samples. *Protist* 157: 31-43.

Thompson, G.A. (2004) Tintinnid diversity trends in the southwestern Atlantic Ocean (29 to 60 degrees S). *Aquatic Microbial Ecology* 35: 93-103.

VanHeukelem L., Lewitus A.J., Kana T.M., Craft N. E. (1994) Improved separations of phytoplankton pigments using temperature-controlled high performance liquid chromatography. *Marine Ecology Progress Series* 114: 303-313.

Woodcock, S., van der Gast, C., Bell, T., Lunn, M., Curtis, T.P., Head, I.M., Sloan, W.T. (2007) Neutral assembly of bacterial communities. *FEMS Microbiology Ecology* 62: 171-180.

Xu, J.P. (2006) Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. *Molecular Ecology* 15: 1713-1731.

Zettler, L.A.A., Gomez, F., Zettler, E., Keenan, B.G., Amils, R., Sogin, M.L. (2002) Eukaryotic diversity in Spain's River of Fire - this ancient and hostile ecosystem hosts a surprising variety of microbial organisms. *Nature* 417: 137-137.